

From: Swope, Sheridan
Sent: Tuesday, November 02, 2004 2:14 PM
To: STIC-Biotech/ChemLib
Subject: 09/940,235

For 09/940,235, pls search:

SID 2, residues 16-383, against the NT and AA data bases.

SID 4, residues 1-106, against the NT and AA data bases.

SID 4, residues 150-259, against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

Scan this
page

(RFE)

T. Port

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:34:19 ; Search time 194.082 Seconds
(without alignments)
1090.970 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901
Sequence: 1 SQLVSVAGTVEGTNQDISL.....IITVMGKRPEGENASYHLA 368

Scoring table: BLOSUM62

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	100.0	440	1 STRP_STREQ	P00779 streptococ
2	1875	98.6	414	2 Q53284	Q53284 streptococ
3	1860	97.8	436	2 Q6UK57	Q6UK57 streptococ
4	1860	97.8	436	2 AAQ73571	AAQ73571 streptoco
5	1859	97.8	440	1 STRP_STRS1	P10519 streptococ
6	1761	92.0	440	2 Q7X0Y1	Q7X0Y1 streptococ
7	1748	92.0	440	2 Q7X0Y7	Q7X0Y7 streptococ
8	1748	92.0	440	2 Q7X0Y8	Q7X0Y8 streptococ
9	1748	92.0	440	2 Q8KSR8	Q8KSR8 streptococ
10	1741	91.6	440	1 STRP_STRPY	P10520 streptococ
11	1741	91.6	440	2 AAP39948	AAP39948 streptoco
12	1734	91.2	440	2 Q7X0Y2	Q7X0Y2 streptococ
13	1697	89.3	440	2 Q7X0Y7	Q7X0Y7 streptococ
14	1682	88.5	440	2 Q8NZA6	Q8NZA6 streptococ
15	1650	86.8	440	2 Q7X0Y3	Q7X0Y3 streptococ
16	1644	86.5	440	2 Q7X0Y0	Q7X0Y0 streptococ
17	1642	86.4	440	2 Q7X0Y5	Q7X0Y5 streptococ
18	1637	86.1	440	2 Q7X0X8	Q7X0X8 streptococ
19	1613	84.9	440	2 Q7X0X9	Q7X0X9 streptococ
20	1604	84.4	440	2 Q7X0X7	Q7X0X7 streptococ
21	1597	84.0	440	2 Q57391	Q57391 streptococ
22	1576	82.9	440	1 STREQ_STRPY	P96471 streptococ
23	1574	82.8	440	2 Q7X0Y6	Q7X0Y6 streptococ
24	1541	81.1	432	2 Q7X0Y4	Q7X0Y4 streptococ
25	728	38.3	141	2 Q7X0X2	Q7X0X2 streptococ
26	727	38.2	141	2 Q7X0X3	Q7X0X3 streptococ
27	724	38.1	141	2 Q7X0W1	Q7X0W1 streptococ
28	723	38.0	141	2 Q7WS87	Q7WS87 streptococ
29	717	37.7	141	2 Q7X0W3	Q7X0W3 streptococ
30	717	37.7	141	2 Q7X0X4	Q7X0X4 streptococ
31	715	37.6	141	2 Q7X0W5	Q7X0W5 streptococ

32	714	37.6	141	2 Q7X0X5	Q7X0X5 streptococ
33	712	37.5	141	2 Q7X0X6	Q7X0X6 streptococ
34	711	37.4	141	2 Q7X0W2	Q7X0W2 streptococ
35	711	37.4	141	2 Q7X0W9	Q7X0W9 streptococ
36	710	37.3	141	2 Q7X0W6	Q7X0W6 streptococ
37	707	37.2	141	2 Q7X0W0	Q7X0W0 streptococ
38	705	37.1	141	2 Q7X0V5	Q7X0V5 streptococ
39	704	37.0	141	2 Q7X0X0	Q7X0X0 streptococ
40	701	36.9	141	2 Q7X0W7	Q7X0W7 streptococ
41	697	36.7	141	2 Q7X0W4	Q7X0W4 streptococ
42	696	36.6	141	2 Q7X0W8	Q7X0W8 streptococ
43	693	36.5	141	2 Q7X0X1	Q7X0X1 streptococ
44	690	36.3	141	2 Q7X0R2	Q7X0R2 streptococ
45	686	36.1	141	2 Q7X0S4	Q7X0S4 streptococ

ALIGNMENTS

RESULT 1

STRP_STREQ STANDARD; PRT; 440 AA.
AC P00779;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DE 01-OCT-2004 (Rel. 45, Last annotation update)
DE Streptokinase C precursor.
GN Name=skc;
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H46A;
RX MEDLINE=85232082; PubMed=2989113;
RA Malke H., Roe B., Ferretti J.J.;
RT "Nucleotide sequence of the streptokinase gene from Streptococcus
RT equisimilis H46A".
RL Gene 34:357-362(1985).
RN [2]
RP SEQUENCE OF 27-440.
RX MEDLINE=83127125; PubMed=6760891;
RA Jackson K.W., Tang J.;
RT "Complete amino acid sequence of streptokinase and its homology with
RT serine proteases".
RL Biochemistry 21:6620-6625(1982).
CC -!- FUNCTION: This protein is not a protease, but it activates
CC plasminogen by complexing with it. As a potential virulence
CC factor, it is thought to prevent the formation of effective fibrin
CC barriers around the site of infection, thereby contributing to the
CC invasiveness of the cells.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K02986; AAA26974.1; --
CC EMBL; X72832; CAAS1351.1; --
CC PIR; A00967; BZSO.
CC PIR; A22801; A22801.
CC PDB; 1L4D; X-ray; B=40-173.
CC PDB; 1L4Z; X-ray; B=26-173.
CC PDB; 1LQK; X-ray; A/B/C/D=177-314.
CC InterPro; IPR004093; Staphylokinase.
CC InterPro; IPR008124; Streptokinase.
CC Pfam; PF02821; Staphylokinase; 3.
CC PRINTS; PR01753; STREPKINASE.
CC 3D-structure; Direct protein sequencing; Plasminogen activation;
KW

Signal; Virulence.
KW SIGNAL 1 26
FT CHAIN 27 440 Streptokinase C.
FT VARIANT 195 195 L -> D.
FT VARIANT 207 200 D -> L.
FT CONFLICT 298 300 EYK -> LEYK (in Ref. 2).
FT CONFLICT 438 438 N -> D (in Ref. 2).
FT STRAND 180 180
FT STRAND 184 194
FT STRAND 184 194
FT TURN 199 200
FT TURN 203 204
FT STRAND 205 205
FT STRAND 209 214
FT TURN 216 217
FT STRAND 219 221
FT STRAND 222 236
FT HELIX 236 239
FT TURN 238 239
FT STRAND 240 252
FT TURN 253 254
FT STRAND 259 260
FT STRAND 268 270
FT STRAND 275 275
FT STRAND 278 280
FT TURN 282 284
FT STRAND 287 289
FT STRAND 292 304
FT TURN 305 306
SQ SEQUENCE 440 AA; 50140 MW; 8FC1F22648ACCT7A CRC64;

Query Match 100.0%; Score 1901; DB 1; Length 440;
Best Local Similarity 100.0%; Pred. No. 4.7e-110;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 42 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 101
QY 61 KADLLKAIQEOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
DB 102 KADLLKAIQEOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 161
QY 121 LLSGHVRVRPYKEKPIQNAQSVVDVEYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 180
DB 162 LLSGHVRVRPYKEKPIQNAQSVVDVEYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 221
QY 181 SOELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
DB 222 SOELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 281
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
DB 282 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
QY 301 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
DB 342 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 401
QY 361 ENASYHLA 368
DB 402 ENASYHLA 409

RESULT 2
Q53284 PRELIMINARY; PRT; 414 AA.
AC Q53284;
DT 01-NOV-1996 (TreeBLrel. 01, Created)
DT 01-NOV-1996 (TreeBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TreeBLrel. 25, Last annotation update)
DE SKC-2.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RA Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
RT Rubiera R., Pedraza A., Padron G., Antuch W., la Fuente J., et al.;
RT "High level expression of streptokinase in Escherichia coli.";
RL Biotechnology 0:1138-1142(1992).
DR EMBL; S46536; AAC60418.1;
DR PDB; 1C4P; X-ray; A/B/C/D=149-285.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
SQ SEQUENCE 414 AA; 47254 MW; F75BE5831B766904 CRC64;
Query Match 98.6%; Score 1875; DB 2; Length 414;
Best Local Similarity 98.6%; Pred. No. 1.8e-108;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 16 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 75
QY 61 KADLLKAIQEOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
DB 76 KADLLKAIQEOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 135
QY 121 LLSGHVRVRPYKEKPIQNAQSVVDVEYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 180
DB 136 LLSGHVRVRPYKEKPIQNAQSVVDVEYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 195
QY 181 SOELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
DB 196 SOELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 255
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
DB 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
QY 301 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
DB 316 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375
QY 361 ENASYHLA 368
DB 376 ENASYHLA 383
RESULT 3
Q6UK57 PRELIMINARY; PRT; 436 AA.
AC Q6UK57;
DT 05-JUL-2004 (TreeBLrel. 27, Created)
DT 05-JUL-2004 (TreeBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TreeBLrel. 27, Last annotation update)
DE Streptokinase.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35666;
RA Costa C.S., Torres F.A.G., Filho S.A.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY368335; AAQ73571.1;
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.

KW Kinase.
SQ SEQUENCE 436 AA; 49658 MW; CA2DD95F957D8F3C CRC64;
Query Match 97.8%; Score 1860; DB 2; Length 436;
Best Local Similarity 97.6%; Pred. No. 1.6e-107; Mismatches 6; Indels 0; Gaps 0;
Matches 359; Conservative 3;
QY 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 38 SOLVSVAGTVEGTNQDISLKFEIDLTSQPAHGKTEQGLSPKSPKPFATDSGAMPHKLE 97
QY 61 KADLLKAIOEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGSVTLPTQVQSF 120
DB 98 KADLLKAIOEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGSVTLPTQVQSF 157
QY 121 LLSGHVVRPYKEKPIQNOAKSVVDVEYVQFTPLNPDHDDFRGLKDTLLKLTLAGDTIT 180
DB 158 LLSGHVVRPYKEKPVQNOAKSVVDVEYVQFTPLNPDHDDFRGLKDTLLKLTLAGDTIT 217
QY 181 SOELLAQAQSILNKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNEQAYRIN 240
DB 218 SOELLAQAQSILNKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNEQAYEIN 277
QY 241 KKSGLNEBINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVNTNELLKSQLLT 300
DB 278 KKSGLNEBINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVNTNELLKSQLLT 337
QY 301 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 360
DB 338 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 397
QY 361 ENASYHLA 368
DB 398 ENASYHLA 405
RESULT 4
AAQ73571 PRELIMINARY; PRT; 436 AA.
AC AAQ73571;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Streptokinase.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus;
OX NCBI_TaxID=119602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35666;
RA Costa C.S., Torres F.A.G., Filho S.A.;
RT "Cloning and expression of a streptokinase gene from Streptococcus
equisimilis (ATCC 35666) in Escherichia coli."
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY368335; AAQ73571.1; --
KW Kinase.
SQ SEQUENCE 436 AA; 49658 MW; CA2DD95F957D8F3C CRC64;
Query Match 97.8%; Score 1860; DB 2; Length 436;
Best Local Similarity 97.6%; Pred. No. 1.6e-107; Mismatches 6; Indels 0; Gaps 0;
Matches 359; Conservative 3;
QY 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 38 SOLVSVAGTVEGTNQDISLKFEIDLTSQPAHGKTEQGLSPKSPKPFATDSGAMPHKLE 97
QY 61 KADLLKAIOEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGSVTLPTQVQSF 120
DB 98 KADLLKAIOEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGSVTLPTQVQSF 157
QY 121 LLSGHVVRPYKEKPIQNOAKSVVDVEYVQFTPLNPDHDDFRGLKDTLLKLTLAGDTIT 180
DB 158 LLSGHVVRPYKEKPVQNOAKSVVDVEYVQFTPLNPDHDDFRGLKDTLLKLTLAGDTIT 217

DB 158 LLSGHVVRPYKEKPVQNOAKSVVDVEYVQFTPLNPDHDDFRGLKDTLLKLTLAGDTIT 217
QY 181 SOELLAQAQSILNKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNEQAYRIN 240
DB 218 SOELLAQAQSILNKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNEQAYEIN 277
QY 241 KKSGLNEBINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVNTNELLKSQLLT 300
DB 278 KKSGLNEBINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVNTNELLKSQLLT 337
QY 301 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 360
DB 338 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 397
QY 361 ENASYHLA 368
DB 398 ENASYHLA 405
RESULT 5
STRP_STRS1 STANDARD; PRT; 440 AA.
AC FL0519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Streptokinase G precursor.
GN Names=kg;
OS Streptococcus sp. (strain 19909).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=69017;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89160265; PubMed=2922269;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
Streptococcus";
RL Nucleic Acids Res. 17:1262-1262 (1989).
CC -!- FUNCTION: This protein is not a protease, but it activates
CC plasminogen by complexing with it. As a potential virulence
CC factor, it is thought to prevent the formation of effective fibrin
CC barriers around the site of infection, thereby contributing to the
CC invasiveness of the cells.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X13400; CAA31766.1; --
DR PIR; S02723; S02723.
DR HSSP; P00779; IQOR.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 Streptokinase G.
SQ SEQUENCE 440 AA; 50199 MW; 5521F8825FE1B6EA CRC64;
Query Match 97.8%; Score 1859; DB 1; Length 440;
Best Local Similarity 97.8%; Pred. No. 1.9e-107; Mismatches 6; Indels 0; Gaps 0;
Matches 360; Conservative 2;
QY 1 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 42 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGKTEQGLSPKSPKPFATDSGAMPHKLE 101

```
QY 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQVQSF 120
DB 102 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQVQSF 161
QY 121 LLSGHVRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDHDDFRPGLKDTKLLKTLAIGDTIT 180
DB 162 LLSGHVRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDHDDFRPGLKDTKLLKTLAIGDTIT 221
QY 181 SOELLAQAQSIILKNKHPGTYIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
DB 222 SOELLAQAQSIILKNKHPGTYIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 281
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
DB 282 KKSGLNEEINNTDLISEKYVVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
QY 301 ASERNLDPRLDYPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPBG 360
DB 342 ASERNLDPRLDYPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPBG 401
QY 361 ENASYHLA 368
DB 402 ENASYHLA 409

RESULT 6
Q7X0Y1
ID Q7X0Y1 PRELIMINARY; PRT; 440 AA.
AC Q7X0Y1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D488;
RX PubMed=14679231;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
RT involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234137; AAP39957.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
KW Kinase.
SQ SEQUENCE 440 AA; 49903 MW; CC4E6E9647043BAC CRC64;

Query Match 92.6%; Score 1761; DB 2; Length 440;
Best Local Similarity 91.6%; Pred. No. 2.3e-101;
Matches 337; Conservative 16; Mismatches 15; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVGVTNODISLKFEEIDLTSPRAGHGKTEQGLSPKSPKFPATDNGAMSHKLE 60
DB 42 SOLVSMAGIVEGTDKKVFINFEEIDLTSPQAHGGKTEQGLSPKSPKFPATDNGAMPHKLE 101
QY 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQVQSF 120
DB 102 KADLLKAIQEQLIANVHNSDGYFEVIDFASDATITDRNGQVYFADKGGSVTLPTQVQSF 161
QY 121 LLSGHVRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDHDDFRPGLKDTKLLKTLAIGDTIT 180
DB 162 LLSGHVRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDHDDFRPGLKDTKLLKTLAIGDTIT 221
QY 181 SOELLAQAQSIILKNKHPGTYIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
DB 222 SOELLAQAQSIILKNKHPGTYIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 281
```

```
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
DB 282 KKSGLNEEINNTDLISEKYVVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
QY 301 ASERNLDPRLDYPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPBG 360
DB 342 ASERNLDPRLDYPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPBG 401
QY 361 ENASYHLA 368
DB 402 AKGSYHLA 409

RESULT 7
Q7W73
ID Q7W73 PRELIMINARY; PRT; 440 AA.
AC Q7W73;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Streptokinase A.
GN OrderedLocusNames=SPs1700;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SSI-1;
RX MEDLINE=22683278; PubMed=12799345;
RA Nakagawa I., Kurokawa K., Yamashita A., Nakata M., Tomiyasu Y.,
RA Okahashi N., Kawabata S., Yamazaki K., Shiba T., Yasunaga T.,
RA Hayashi H., Hattori M., Hamada S.;
RT "Genome sequence of an M3 strain of Streptococcus pyogenes reveals a
RT large-scale genomic rearrangement in invasive strains and new insights
RT into phage evolution.";
RL Genome Res. 13:1042-1055(2003).
DR EMBL; AP005146; BAC64795.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKINASE.
KW Kinase.
SQ SEQUENCE 440 AA; 49911 MW; 7CC44F4026E3975 CRC64;

Query Match 92.0%; Score 1748; DB 2; Length 440;
Best Local Similarity 91.3%; Pred. No. 1.5e-100;
Matches 336; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVGVTNODISLKFEEIDLTSPRAGHGKTEQGLSPKSPKFPATDNGAMSHKLE 60
DB 42 SOLVSMAGIVEGTDKKVFINFEEIDLTSPQAHGGKTEQGLSPKSPKFPATDNGAMPHKLE 101
QY 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQVQSF 120
DB 102 KADLLKAIQEQLIANVHNSDGYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQVQSF 161
QY 121 LLSGHVRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDHDDFRPGLKDTKLLKTLAIGDTIT 180
DB 162 LLSGHVRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDHDDFRPGLKDTKLLKTLAIGDTIT 221
QY 181 SOELLAQAQSIILKNKHPGTYIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
DB 222 SOELLAQAQSIILKNKHPGTYIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 281
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
DB 282 KKSGLNEEINNTDLISEKYVVLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
QY 301 ASERNLDPRLDYPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPBG 360
```

Db 342 ASERNLDFRDLVDPDCAKALLNNLDADFIMDYTLTGKVEDNHDKNRIVTVYMGKRPKG 401
Qy 361 ENASYHLA 368
|
|
|
|
Db 402 AKGSYHLA 409

RESULT 8

ID Q7X0Y8 PRELIMINARY; PRT; 440 AA.
AC Q7X0Y8;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=88-019;
RX PubMed=14679231;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY234129; AAP39949.1; -;
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
KW Kinase.
SQ SEQUENCE 440 AA; 49885 MW; 69DB44F4026E3975 CRC64;

Query Match 92.0%; Score 1748; DB 2; Length 440;

Best Local Similarity 91.3%; Pred. No. 1.5e-100;

Matches 336; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVGVTGNTQDLSKPFEDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 60
|
|
|
|
Db 42 SOLVSVAGIVEGTDKKVFNFEIDLTSQAHGKTEQGLSPKSPFATDNGAMPHKLE 101
|
|
|
|
Qy 61 KADLLKAIQOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGSVTLTPQVQEF 120
|
|
|
|
Db 102 KADLLKAIQOLIANVHNSDGYFEVIDFASDATITDRNGKVYFADKGSVTLTPQVQEF 161
|
|
|
|
Qy 121 LLSGHRVVRPYKEKPYQNAKSDVVEYTVQFTPLNPDFFRPGDKTKLLKTLAIGDTIT 180
|
|
|
|
Db 162 LLSGHRVVRPYKEKPYQNAKSDVVKYTVQFTPLNPDFFRPGDKTKLLKTLAIGDTIT 221
|
|
|
|
Qy 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIFRTILPMDQEFYRKNREAYRIN 240
|
|
|
|
Db 222 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIFRTILPMDQEFYRKNREAYGIN 281
|
|
|
|
Qy 241 KKSGLNEEINNTDLISEKYVVKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
|
|
|
|
Db 282 KKSGLNEEINNTDLISEKYVVKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
|
|
|
|
Qy 301 ASERNLDFRDLVDPDCAKALLNNLDADFIMDYTLTGKVEDNHDKNRIVTVYMGKRPKG 360
|
|
|
|
Db 342 ASERNLDFRDLVDPDCAKALLNNLDADFIMDYTLTGKVEDNHDKNRIVTVYMGKRPKG 401
|
|
|
|
Qy 361 ENASYHLA 368
|
|
|
|
Db 402 AKGSYHLA 409

RESULT 9

Q8K5R8 PRELIMINARY; PRT; 440 AA.
ID Q8K5R8
AC Q8K5R8;

DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Streptokinase A.
GN Name=ska; OrderedLocusNames=SpyM3_1698;
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
RL EMBL; AE014169; AAM80305.1; -;
DR HSSP; Q53284; IC4P.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasmaenogen activator activity; IEA.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR InterPro; IPR008124; Streptokinase.
DR Pfam; PF02821; Staphylokinase; 3.
DR PRINTS; PR01753; STREPKNINASE.
DR Complete proteome; Kinase.
KW Complete proteome; Kinase.
SQ SEQUENCE 440 AA; 49911 MW; 7CCB44F4026E3975 CRC64;

Query Match 92.0%; Score 1748; DB 2; Length 440;
Best Local Similarity 91.3%; Pred. No. 1.5e-100;
Matches 336; Conservative 15; Mismatches 17; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVGVTGNTQDLSKPFEDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 60
|
|
|
|
Db 42 SOLVSVAGIVEGTDKKVFNFEIDLTSQAHGKTEQGLSPKSPFATDNGAMPHKLE 101
|
|
|
|
Qy 61 KADLLKAIQOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKGSVTLTPQVQEF 120
|
|
|
|
Db 102 KADLLKAIQOLIANVHNSDGYFEVIDFASDATITDRNGKVYFADKGSVTLTPQVQEF 161
|
|
|
|
Qy 121 LLSGHRVVRPYKEKPYQNAKSDVVEYTVQFTPLNPDFFRPGDKTKLLKTLAIGDTIT 180
|
|
|
|
Db 162 LLSGHRVVRPYKEKPYQNAKSDVVKYTVQFTPLNPDFFRPGDKTKLLKTLAIGDTIT 221
|
|
|
|
Qy 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIFRTILPMDQEFYRKNREAYRIN 240
|
|
|
|
Db 222 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIFRTILPMDQEFYRKNREAYGIN 281
|
|
|
|
Qy 241 KKSGLNEEINNTDLISEKYVVKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
|
|
|
|
Db 282 KKSGLNEEINNTDLISEKYVVKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
|
|
|
|
Qy 301 ASERNLDFRDLVDPDCAKALLNNLDADFIMDYTLTGKVEDNHDKNRIVTVYMGKRPKG 360
|
|
|
|
Db 342 ASERNLDFRDLVDPDCAKALLNNLDADFIMDYTLTGKVEDNHDKNRIVTVYMGKRPKG 401
|
|
|
|
Qy 361 ENASYHLA 368
|
|
|
|
Db 402 AKGSYHLA 409

RESULT 10
STRP_STRPY STANDARD; PRT; 440 AA.
ID STRP_STRPY
AC P10520;
DT 01-JUL-1989 (Rel. 11, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)


```
RESULT 12
Q7X0Y2
ID Q7X0Y2 PRELIMINARY; PRT; 440 AA.
AC Q7X0Y2;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D306;
RX PubMed=14679231;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY2341136; AAP39956.1; -
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
KW Kinase.
SQ SEQUENCE 440 AA; 50188 MW; 4CB42AC2A9062C2E CRC64;
Query Match 91.2%; Score 1734; DB 2; Length 440;
Best Local Similarity 90.2%; Pred. No. 1.1e-99;
Matches 332; Conservative 18; Mismatches 18; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVEGTNQDLSLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 42 SOLVSMAGIVEGTDKKVFINFPEIDLTSQPAHGGKTEQGLSPKSPFATDSAMPHKLE 101
Qy 61 KADLLKAIQOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
Db 102 KADLLKAIQOLIANVHNSDGYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 161
Qy 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFKGLKDTLLKTLAIGDTIT 180
Db 162 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFKGLKDTLLKTLAIGDTIT 221
Qy 181 SOELLAQAQSILNKHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVNREQAYRIN 240
Db 222 SOELLAQAQSILNESHDPDYTIYERDSSIVTHDNDIFRTILPMDQEFYRVNREQAYRIN 281
Qy 241 ASERNLDFRDLDPDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHNTNRIITVYMGKRPEG 300
Db 342 ASERNLDFRDLDPDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHNTNRIITVYMGKRPEG 401
Qy 361 ENASYHLA 368
Db 402 AKGSYHLA 409
RESULT 13
Q7X0Y7
ID Q7X0Y7 PRELIMINARY; PRT; 440 AA.
AC Q7X0Y7;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Streptokinase.
GN Name=ska;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D306;
RX PubMed=14679231;
RA Kalia A., Bessen D.E.;
RT "Natural selection and evolution of streptococcal virulence genes
involved in tissue-specific adaptations.";
RL J. Bacteriol. 186:110-121(2004).
DR EMBL; AY2341136; AAP39956.1; -
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0008243; F:plasminogen activator activity; IEA.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
KW Kinase.
SQ SEQUENCE 440 AA; 50160 MW; 1C5AF07907EC7AC8 CRC64;
Query Match 89.3%; Score 1697; DB 2; Length 440;
Best Local Similarity 88.6%; Pred. No. 2.2e-97;
Matches 326; Conservative 19; Mismatches 23; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVEGTNQDLSLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 42 SOLVSMAGIVEGTDKKVFINFPEIDLTSQPAHGGKTEQGLSPKSPFATNSSAMPHKLE 101
Qy 61 KADLLKAIQOLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
Db 102 KADLLKAIQOLIANVHNSDGYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 161
Qy 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFKGLKDTLLKTLAIGDTIT 180
Db 162 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFKGLKDTLLKTLAIGDTIT 221
Qy 181 SOELLAQAQSILNKHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVNREQAYRIN 240
Db 222 SOELLAQAQSILNESHDPDYTIYERDSSIVTHDNDIFRTILPMDQEFYRVNREQAYRIN 281
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 282 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 341
Qy 301 ASERNLDFRDLDPDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHNTNRIITVYMGKRPEG 360
Db 342 ASERNLDFRDLDPDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHNTNRIITVYMGKRPEG 401
Qy 361 ENASYHLA 368
Db 402 ENASYHLA 409
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:42:25 ; Search time 44.7397 Seconds
(without alignments)
545.489 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901
Sequence: 1 SQLVSVAGTVEGTNQDISL.....ITVYMKRPEGENASYHLA 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1901	100.0	414	3	US-09-211-542A-6
2	1901	100.0	414	6	5240845-1
3	1901	100.0	440	2	US-08-560-098A-52
4	1901	100.0	795	3	US-09-211-542A-2
5	1883.5	99.1	800	2	US-08-488-940-4
6	1883.5	99.1	813	2	US-08-488-940-3
7	1883.5	99.1	1181	2	US-08-488-940-2
8	1883.5	99.1	1194	2	US-08-488-940-1
9	1883.5	99.1	1194	2	US-08-488-940-17
10	1883.5	99.1	1194	2	US-08-488-940-18
11	1875	98.6	369	1	US-07-854-596B-31
12	1875	98.6	401	3	US-09-374-038-1
13	1875	98.6	401	4	US-09-658-179-1
14	1875	98.6	413	3	US-09-374-038-2
15	1875	98.6	413	4	US-09-658-179-2
16	1875	98.6	415	1	US-07-854-596B-26
17	1875	98.6	435	1	US-07-854-596B-19
18	1875	98.6	440	1	US-07-854-596B-15
19	1875	98.6	483	1	US-07-854-596B-43
20	1875	98.6	483	1	US-07-854-596B-47
21	1875	98.6	499	1	US-07-854-596B-28
22	1875	98.6	747	1	US-07-854-596B-40
23	1875	98.6	859	1	US-07-854-596B-35
24	1845	97.1	413	2	US-08-759-599-12
25	1845	97.1	413	3	US-09-294-457-12
26	1819	95.7	384	3	US-09-374-038-4
27	1819	95.7	384	4	US-09-658-179-4

28	1815	95.5	372	3	US-09-374-038-3	Sequence 3, Appli
29	1815	95.5	372	4	US-09-658-179-3	Sequence 3, Appli
30	1693.5	89.1	736	3	US-09-211-542A-4	Sequence 4, Appli
31	1686	88.7	356	3	US-09-211-542A-12	Sequence 12, Appl
32	785	41.3	150	3	US-09-211-542A-14	Sequence 14, Appl
33	275	14.5	286	4	US-09-078-091-4	Sequence 4, Appli
34	267	14.0	286	4	US-09-078-091-2	Sequence 2, Appli
35	118	6.2	442	4	US-09-328-352-5748	Sequence 5748, Ap
36	115.5	6.1	1289	2	US-08-853-659A-51	Sequence 51, Appl
37	115	6.0	21	2	US-08-759-599-7	Sequence 7, Appli
38	115	6.0	21	3	US-09-294-457-7	Sequence 7, Appli
39	111	5.8	1588	5	PCT-US93-07261-11	Sequence 11, Appl
40	111	5.8	1663	5	PCT-US93-07261-16	Sequence 16, Appl
41	109.5	5.8	924	4	US-09-248-796A-18798	Sequence 18798, A
42	109.5	5.8	1220	2	US-08-680-326-38	Sequence 38, Appl
43	108.5	5.7	585	4	US-09-107-532A-5494	Sequence 5494, Ap
44	108	5.7	1183	2	US-08-447-031A-2	Sequence 2, Appli
45	106	5.6	1833	4	US-08-621-944A-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-211-542A-6
; Sequence 6, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpei, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-211-542A-6

Query Match 100.0%; Score 1901; DB 3; Length 414;
Best Local Similarity 100.0%; Pred. No. 4.6e-169;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SQLVSVAGTVEGTNQDISLKFELDITSRPAHGKTEQGLSPKSPKPFATDSGAMSHKLE 60
Db 16 SQLVSVAGTVEGTNQDISLKFELDITSRPAHGKTEQGLSPKSPKPFATDSGAMSHKLE 75

Qy 61 KADLLKAIQEOQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
Db |||||
Qy 76 KADLLKAIQEOQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 135
Db |||||
Qy 121 LLSGHVVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTKLLKTLAIGDTIT 180
Db |||||
Qy 136 LLSGHVVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTKLLKTLAIGDTIT 195
Db |||||
Qy 181 SOELLAQAQSIILKNHHPGYTIYERDSSIVTHNDIFRTILPMDQEFYVKNREQAYRIN 240
Db |||||
Qy 196 SOELLAQAQSIILKNHHPGYTIYERDSSIVTHNDIFRTILPMDQEFYVKNREQAYRIN 255
Db |||||
Qy 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db |||||
Qy 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Db |||||
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db |||||
Qy 316 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375
Db |||||
Qy 361 ENASYHLA 368
Db |||||
Qy 376 ENASYHLA 383
Db |||||

RESULT 2

5240845-1
; Patent No. 5240845
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO; KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS
; NUMBER OF SEQUENCES: 65
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/549,049
; FILING DATE: 06-JUL-1990
; SEQ ID NO:1
; LENGTH: 414
5240845-1

Query Match 100.0%; Score 1901; DB 6; Length 414;
Best Local Similarity 100.0%; Pred. No. 4.6e-169;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 60
Db |||||
Qy 16 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 75
Db |||||
Qy 61 KADLLKAIQEOQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
Db |||||
Qy 76 KADLLKAIQEOQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 135
Db |||||
Qy 121 LLSGHVVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTKLLKTLAIGDTIT 180
Db |||||
Qy 136 LLSGHVVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTKLLKTLAIGDTIT 195
Db |||||
Qy 181 SOELLAQAQSIILKNHHPGYTIYERDSSIVTHNDIFRTILPMDQEFYVKNREQAYRIN 240
Db |||||
Qy 196 SOELLAQAQSIILKNHHPGYTIYERDSSIVTHNDIFRTILPMDQEFYVKNREQAYRIN 255
Db |||||
Qy 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db |||||
Qy 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Db |||||
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db |||||
Qy 316 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375
Db |||||
Qy 361 ENASYHLA 368
Db |||||
Qy 376 ENASYHLA 383
Db |||||

RESULT 3
US-08-560-098A-52
; Sequence 52, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA: P 44 40 892.7
; APPLICATION NUMBER: 17-NOV-1994
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-52

Query Match 100.0%; Score 1901; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 5.1e-169;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 60
Db |||||
Qy 42 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 101
Db |||||
Qy 61 KADLLKAIQEOQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
Db |||||
Qy 102 KADLLKAIQEOQLIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 161
Db |||||
Qy 121 LLSGHVVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTKLLKTLAIGDTIT 180
Db |||||
Qy 162 LLSGHVVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTKLLKTLAIGDTIT 221
Db |||||
Qy 181 SOELLAQAQSIILKNHHPGYTIYERDSSIVTHNDIFRTILPMDQEFYVKNREQAYRIN 240
Db |||||
Qy 222 SOELLAQAQSIILKNHHPGYTIYERDSSIVTHNDIFRTILPMDQEFYVKNREQAYRIN 281
Db |||||
Qy 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db |||||
Qy 282 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 341
Db |||||
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db |||||
Qy 342 ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 401
Db |||||

QY 361 ENASYHLA 368
| | | | |
DB 402 ENASYHLA 409

RESULT 4

US-09-211-542A-2
; Sequence 2, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)443-9292
; TELEFAX: (617)443-0004
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 795 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-211-542A-2

Query Match 100.0%; Score 1901; DB 3; Length 795;
Best Local Similarity 100.0%; Pred. No. 1.3e-168;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SOLVSVAGTVEGTNQDISLKPFIDLTSPRPAHGKTEQGLSPKSPKPFATDGSAMSHKLE 60
DB 397 SOLVSVAGTVEGTNQDISLKPFIDLTSPRPAHGKTEQGLSPKSPKPFATDGSAMSHKLE 456
QY 61 KADLLKAIQEQLIANVHSDNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 120
DB 457 KADLLKAIQEQLIANVHSDNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 516
QY 121 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 180
DB 517 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 576
QY 181 SQELLAQAQSILKNKHPGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
DB 577 SQELLAQAQSILKNKHPGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 636
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
DB 637 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 696
QY 301 ASERNLDFRDLYDRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVTMGRPEG 360
| | | | |

DB 697 ASERNLDFRDLYDRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVTMGRPEG 756
QY 361 ENASYHLA 368
| | | | |
DB 757 ENASYHLA 764

RESULT 5

US-08-488-940-4
; Sequence 4, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-4

Query Match 99.1%; Score 1883.5; DB 2; Length 800;
Best Local Similarity 99.7%; Pred. No. 5.5e-167;
Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 SOLVSVAGTVEGTNQDISLKPFIDLTSPRPAHGKTEQGLSPKSPKPFATDGSAMSHKLE 60
DB 3 SOLVSVAGTVEGTNQDISLKPFIDLTSPRPAHGKTEQGLSPKSPKPFATDGSAMSHKLE 62
QY 61 KADLLKAIQEQLIANVHSDNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 120
DB 63 KADLLKAIQEQLIANVHSDNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 122
QY 121 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 180
DB 123 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 181
QY 181 SQELLAQAQSILKNKHPGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
DB 182 SQELLAQAQSILKNKHPGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 241
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
DB 242 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 301
QY 301 ASERNLDFRDLYDRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDHDTNRIITVTMGRPEG 360
| | | | |

Db 302 ASERNLDFRDLDPDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNTRIIITVYMGKPEG 361
QY 361 ENASYHLA 368
Db 362 ENASYHLA 369

RESULT 6

US-08-940-3
; Sequence 3, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 813 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-3

Query Match 99.1%; Score 1883.5; DB 2; Length 813;
Best Local Similarity 99.7%; Pred. No. 5.6e-167;
Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 SOLVSVAGTVGVTGQDLSLKFFEDLTSRPAHGKTEOGLSPKSPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVGVTGQDLSLKFFEDLTSRPAHGKTEOGLSPKSPFATDSGAMSHKLE 75
QY 61 KADLLKAIQEQLIANVHSNDDYFEVIDFASDATITDRNGKVYFADKXGSVTLPTQPVQEF 120
Db 76 KADLLKAIQEQLIANVHSNDDYFEVIDFASDATITDRNGKVYFADKXGSVTLPTQPVQEF 135
QY 121 LLSGHVRVRPYKEPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 180
Db 136 LLSGHVRVR - YKEKPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 194
QY 181 SOELLAQAQSIILKNKHPGTYIYERDSSIYTHNDIFRTILPMDQEFYRVKNEQAYRIN 240
Db 195 SOELLAQAQSIILKNKHPGTYIYERDSSIYTHNDIFRTILPMDQEFYRVKNEQAYRIN 254
QY 241 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDFDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 255 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDFDRSHLKLFTIKYVDVDTNELLKSQLLT 314

QY 301 ASERNLDFRDLDPDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNTRIIITVYMGKPEG 360
Db 315 ASERNLDFRDLDPDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNTRIIITVYMGKPEG 374
QY 361 ENASYHLA 368
Db 375 ENASYHLA 382

RESULT 7

US-08-488-940-2
; Sequence 2, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1181 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-2

Query Match 99.1%; Score 1883.5; DB 2; Length 1181;
Best Local Similarity 99.7%; Pred. No. 9.9e-167;
Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 SOLVSVAGTVGVTGQDLSLKFFEDLTSRPAHGKTEOGLSPKSPFATDSGAMSHKLE 60
Db 384 SOLVSVAGTVGVTGQDLSLKFFEDLTSRPAHGKTEOGLSPKSPFATDSGAMSHKLE 443
QY 61 KADLLKAIQEQLIANVHSNDDYFEVIDFASDATITDRNGKVYFADKXGSVTLPTQPVQEF 120
Db 444 KADLLKAIQEQLIANVHSNDDYFEVIDFASDATITDRNGKVYFADKXGSVTLPTQPVQEF 503
QY 121 LLSGHVRVRPYKEPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 180
Db 504 LLSGHVRVR - YKEKPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 562
QY 181 SOELLAQAQSIILKNKHPGTYIYERDSSIYTHNDIFRTILPMDQEFYRVKNEQAYRIN 240
Db 563 SOELLAQAQSIILKNKHPGTYIYERDSSIYTHNDIFRTILPMDQEFYRVKNEQAYRIN 622
QY 241 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDFDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 623 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDFDRSHLKLFTIKYVDVDTNELLKSQLLT 682

636	DB	KSGLNEENNTDLSSEKYVYLKKGEPYDPDRSHLKLFTIKYVOVDVTNELLKSQLLT	695
301	QY	ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDDTNRILITVMGKRPEG	360
696	DB	ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDDTNRILITVMGKRPEG	755
361	QY	ENASYHLA 368	
756	DB	ENASYHLA 763	

IIS-08-48

```

RESULT 9
US-08-488-940-17
; Sequence 17, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Read, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-17

```

Query Match	99.1%;	Score 1883.5;	DB 2;	Length 1194;
Best Local Similarity	99.7%;	Pred. No. 1e-166;		
Matches 367;	Conservative	0;	Mismatches 0;	Indels 1; Gaps 1;
Qy	1	SQLVVSVAGTVTEGNTQDILSKFFPEIDLTSPAHGGKTEQGLSPKSKPPATDSGAMSHKLE	60	
Db	797	SQLVVSVAGTVTEGNTQDILSKFFPEIDLTSPAHGGKTEQGLSPKSKPPATDSGAMSHKLE	856	
Qy	61	KADLLKAIQQLANVHSNDYFEVIDFASDATITDRNGKVFYFADKOGSVTLTPQVQEF	120	
Db	857	KADLLKAIQQLANVHSNDYFEVIDFASDATITDRNGKVFYFADKOGSVTLTPQVQEF	916	
Qy	121	LLSGHVRVRPYKPKPIQNAQSVVDVEYVQFTPLNPDDDFRPLGKDTKLLKTLAIGDTIT	180	
Db	917	LLSGHVRVR-YKPKPIQNAQSVVDVEYVQFTPLNPDDDFRPLGKDTKLLKTLAIGDTIT	975	
Qy	181	SOELLAAQASILANKNHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYTVKRNROQAVRIN	240	
Db	976	SOELLAAQASILANKNHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYTVKRNROQAVRIN	1035	
Qy	241	KKSGLNEBEIWNNTDLISSEKYTVLVKKGEKPYDPDFDRSHLKLPTIKVYDVTNELLKSEQLLT	300	

Db 1036 KKSGLNEEINNTDLISEKYVYLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 1095
QY 301 ASERNLDFRDLDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKPEG 360
Db 1096 ASERNLDFRDLDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKPEG 1155
QY 361 ENASYHLA 368
Db 1156 ENASYHLA 1163

RESULT 10
US-08-488-940-18
; Sequence 18, Application US/08488940
; Patent No. 5854049
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: PLASMIN-RESISTANT STREPTOKINASE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,940
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 05433/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1194 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-940-18

Query Match 99.1%; Score 1883.5; DB 2; Length 1194;
Best Local Similarity 99.7%; Pred. No. 1e-166;
Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGKTEQGLSPKSPFPATDSGAMSHKLE 60
Db 797 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGKTEQGLSPKSPFPATDSGAMSHKLE 856
QY 61 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 120
Db 857 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 916
QY 121 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 180
Db 917 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 975
QY 181 SQELLAQAQSILNKNHPGTYIYERDSSIVTHDNDIFRTILPMDQEFYTYRKNREQAYRIN 240
Db 976 SQELLAQAQSILNKNHPGTYIYERDSSIVTHDNDIFRTILPMDQEFYTYRKNREQAYRIN 1035

QY 241 KKSGLNEEINNTDLISEKYVYLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 1036 KKSGLNEEINNTDLISEKYVYLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 1095
QY 301 ASERNLDFRDLDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKPEG 360
Db 1096 ASERNLDFRDLDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKPEG 1155
QY 361 ENASYHLA 368
Db 1156 ENASYHLA 1163

RESULT 11
US-07-854-596B-31
; Sequence 31, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-854-596B-31

Query Match 98.6%; Score 1875; DB 1; Length 369;
Best Local Similarity 98.8%; Pred. No. 1e-166;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGKTEQGLSPKSPFPATDSGAMSHKLE 60
Db 2 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGKTEQGLSPKSPFPATDSGAMSHKLE 61
QY 61 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 120
Db 62 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 121
QY 121 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 180
Db 122 LLSGHRVRPYKEKPIQONAKSVDEVYTVQFTPLNPDDDFRPGDKTKLLKTLAIGDTIT 181
QY 181 SQELLAQAQSILNKNHPGTYIYERDSSIVTHDNDIFRTILPMDQEFYTYRKNREQAYRIN 240

Db 182 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 241
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 242 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 301
Qy 301 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db 302 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 361
Qy 361 ENASYHLA 368
Db 362 ENASYHLA 369

RESULT 12
US-09-374-038-1
; Sequence 1, Application US/09374038
; Patent No. 6309873
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6309873
; CURRENT APPLICATION NUMBER: US/09/374,038
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-374-038-1

Query Match 98.6%; Score 1875; DB 3; Length 401;
Best Local Similarity 98.6%; Pred. No. 1.2e-166;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 3 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMPHKLE 62
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLTPQVQEF 120
Db 63 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLTPQVQEF 122
Qy 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDDFRPGLKDTLLKTLAIGDTIT 180
Db 123 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDDFRPGLKDTLLKTLAIGDTIT 182
Qy 181 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 240
Db 183 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 242
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 243 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 302
Qy 301 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db 303 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 362
Qy 361 ENASYHLA 368
Db 363 ENASYHLA 370

Query Match 98.6%; Score 1875; DB 3; Length 401;
Best Local Similarity 98.6%; Pred. No. 1.2e-166;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 3 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMPHKLE 62
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLTPQVQEF 120
Db 63 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLTPQVQEF 122
Qy 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDDFRPGLKDTLLKTLAIGDTIT 180
Db 123 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDDFRPGLKDTLLKTLAIGDTIT 182
Qy 181 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 240
Db 183 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 242
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 243 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 302
Qy 301 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db 303 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 362
Qy 361 ENASYHLA 368
Db 363 ENASYHLA 370

RESULT 13
US-09-658-179-1
; Sequence 1, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-658-179-1

Query Match 98.6%; Score 1875; DB 4; Length 401;
Best Local Similarity 98.6%; Pred. No. 1.2e-166;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 3 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMPHKLE 62
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLTPQVQEF 120
Db 63 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLTPQVQEF 122
Qy 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDDFRPGLKDTLLKTLAIGDTIT 180
Db 123 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDDFRPGLKDTLLKTLAIGDTIT 182
Qy 181 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 240
Db 183 SQELLAQAQSILNKHGTHGTYIYERDSSIVTHDNDIPRTILPMDQEPYTHVKNREQAYEIN 242
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
Db 243 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 302
Qy 301 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db 303 ASERNLDFRDLDPDRKAKLLNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 362
Qy 361 ENASYHLA 368
Db 363 ENASYHLA 370

RESULT 14
US-09-374-038-2
; Sequence 2, Application US/09374038
; Patent No. 6309873
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6309873

; CURRENT APPLICATION NUMBER: US/09/374,038
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-374-038-2

Query Match 98.6%; Score 1875; DB 3; Length 413;
Best Local Similarity 98.6%; Pred. No. 1.2e-166;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	SQVSVAGTVEGTNQDISLKFFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMPHKLE	60
Db	3	SQVSVAGTVEGTNQDISLKFFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMPHKLE	62
Qy	61	KADLLKAIQEOLIANVHSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF	120
Db	63	KADLLKAIQEOLIANVHSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF	122
Qy	121	LLSGHVRVPYKEKPIQNOAKSVDEYTVQFTPLNPDDDFRPGDKOTKLLKTLAIGDTIT	180
Db	123	LLSGHVRVPYKEKPIQNOAKSVDEYTVQFTPLNPDDDFRPGDKOTKLLKTLAIGDTIT	182
Qy	181	SOELLAAQOSILNKHPGYTIYERDSSIVTHNDIFRTPILPMDQEFYVKNREQAYRIN	240
Db	183	SOELLAQAQOSILNKHPGYTIYERDSSIVTHNDIFRTPILPMDQEFYVKNREQAYEIN	242
Qy	241	KKSGLINEEINNTDLISEKYVVLKKEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT	300
Db	243	KKSGLINEEINNTDLISEKYVVLKKEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT	302
Qy	301	ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG	360
Db	303	ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG	362
Qy	361	ENASYHLA 368	
Db	363	ENASYHLA 370	

RESULT 15
US-09-658-179-2
; Sequence 2, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-658-179-2

Query Match 98.6%; Score 1875; DB 4; Length 413;
Best Local Similarity 98.6%; Pred. No. 1.2e-166;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy	1	SQVSVAGTVEGTNQDISLKFFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMPHKLE	60
----	---	--	----

Db	3	SQVSVAGTVEGTNQDISLKFFEDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMPHKLE	62
Qy	61	KADLLKAIQEOLIANVHSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF	120
Db	63	KADLLKAIQEOLIANVHSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF	122
Qy	121	LLSGHVRVPYKEKPIQNOAKSVDEYTVQFTPLNPDDDFRPGDKOTKLLKTLAIGDTIT	180
Db	123	LLSGHVRVPYKEKPIQNOAKSVDEYTVQFTPLNPDDDFRPGDKOTKLLKTLAIGDTIT	182
Qy	181	SOELLAQAQOSILNKHPGYTIYERDSSIVTHNDIFRTPILPMDQEFYVKNREQAYRIN	240
Db	183	SOELLAQAQOSILNKHPGYTIYERDSSIVTHNDIFRTPILPMDQEFYVKNREQAYEIN	242
Qy	241	KKSGLINEEINNTDLISEKYVVLKKEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT	300
Db	243	KKSGLINEEINNTDLISEKYVVLKKEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT	302
Qy	301	ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG	360
Db	303	ASERNLDFRDLYDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG	362
Qy	361	ENASYHLA 368	
Db	363	ENASYHLA 370	

Search completed: November 3, 2004, 23:56:58
Job time : 45.7397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2004, 00:02:46 ; Search time 120.986 Seconds
(without alignments)
2161.983 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901

Sequence: 1 SQLVSVAGTVEGTNQDISL.....IITYMGRPEGENASYHLA 368

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+P2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US0940235/runat_03112004_174039_11280/app_query.fasta_1.1045
-DB=Issued Patents NA -QFT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0940235@cgn_1_177@runat_03112004_174039_11280 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBL=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:
- 6: /cgn2_6/ptodata/1/ina/backfileseq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	100.0	1242	2	US-08-568-393B-1
2	1901	100.0	1242	3	US-09-211-542A-5
3	1901	100.0	1242	6	5240845-2
4	1901	100.0	1262	3	5240845-3
5	1901	100.0	2385	3	US-09-211-542A-1
6	1897	99.8	1242	2	US-08-568-393B-2
7	1875	98.6	1119	1	US-07-854-596B-30
8	1875	98.6	1209	3	US-09-374-038-11
9	1875	98.6	1209	4	US-09-658-179-11
10	1875	98.6	1245	1	US-07-703-778D-1
11	1875	98.6	1245	3	US-09-374-038-12
12	1875	98.6	1245	4	US-09-658-179-12

13	1875	98.6	1257	1	US-07-854-596B-25
14	1875	98.6	1317	1	US-07-854-596B-18
15	1875	98.6	1335	1	US-07-854-596B-14
16	1875	98.6	1458	1	US-07-854-596B-42
17	1875	98.6	1457	1	US-07-854-596B-46
18	1875	98.6	1512	1	US-07-854-596B-27
19	1875	98.6	2253	1	US-07-854-596B-39
20	1875	98.6	2589	1	US-07-854-596B-34
21	1871.5	98.4	2566	3	US-08-488-940-19
22	1819	95.7	1158	3	US-09-374-038-14
23	1819	95.7	1158	4	US-09-658-179-13
24	1815	95.5	1122	3	US-09-374-038-13
25	1815	95.5	1122	4	US-09-658-179-13
26	1693.5	89.1	2208	3	US-09-211-542A-3
27	1686	88.7	1068	3	US-09-211-542A-11
28	785	41.3	450	3	US-09-211-542A-13
29	275	14.5	1181	4	US-09-078-091-3
30	267	14.0	1180	4	US-09-078-091-1
31	118	6.2	1329	4	US-09-328-352-1622
32	115.5	6.1	3870	2	US-08-853-659A-26
33	115.5	6.1	8967	2	US-08-853-659A-6
34	115.5	6.1	8967	2	US-08-853-659A-9
35	115.5	6.1	8967	2	US-08-853-659A-64
36	115.5	6.1	8967	2	US-08-853-659A-67
37	115.5	6.1	24701	2	US-08-853-659A-2
38	115.5	6.1	24701	2	US-08-853-659A-3
39	115.5	6.1	24701	2	US-08-853-659A-60
40	115.5	6.1	24701	2	US-08-853-659A-61
41	112.5	5.9	1664976	4	US-08-916-421B-1
42	112.5	5.9	1664976	4	US-09-692-570-1
43	111	5.8	4766	5	PCT-US93-07261-10
44	110	5.8	3017	4	US-09-799-451-153
45	109.5	5.8	2775	4	US-09-248-796A-4695

ALIGNMENTS

RESULT 1

US-08-568-393B-1
; Sequence 1, Application US/08568393B
; Patent No. 5876999
; GENERAL INFORMATION:
; APPLICANT: Hua-Lin Wu
; APPLICANT: Guey-Yueh Shi
; TITLE OF INVENTION: Preparation of novel streptokinase
; TITLE OF INVENTION: mutants as improved thrombolytic agents
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jeing & Chang
; STREET: Two No. 5876999th Second Street, Suite 290
; CITY: San Jose
; STATE: California
; COUNTRY: USA
; ZIP: 95113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
; MEDIUM TYPE: storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1 on Window 3.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,393B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chi-Ping Chang
; REGISTRATION NUMBER: 37,798
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (408) 288-8585
; TELEFAX: (408) 288-8386
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

LENGTH: 1242 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL: N
ANTI-SENSE: N
ORIGINAL SOURCE:
ORGANISM: Streptococcus equisimilis H46A
INDIVIDUAL ISOLATE: Maile, H., Roe, B., and Ferretti, J. J.;
INDIVIDUAL ISOLATE: "Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A" from Gene 34:357-362 (1985).
CELL TYPE: Streptococcus equisimilis H46A
US-09-568-393B-1

Alignment Scores:
Pred. No.: 2,49e-203 Length: 1242
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-08-568-393B-1 (1-1242)

Qy	1	SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu	20
Db	46	AGCCAAATAGTTGTTAGCGTTCTGCTACTGTGTAGGGGACGAATCAAGACATTAAGTCTT	105
Qy	21	LysPheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly	40
Db	106	AAATTTTGAATCGAATCAATCAACACCTGCTCATGGAGGAAGACAGACGACGAGC	165
Qy	41	LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu	60
Db	166	TTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTGCATAACTTGA	225
Qy	61	LysAlaAspLeuLeuLysAlaIleGlnGluLeuLeuAlaAsnValHisSerAsnAsp	80
Db	226	AAAGCTGACTTAAAGGCTATTCAAGAACCAATTTGATCGCTAAACGTCCACAGTAACGAC	285
Qy	81	AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys	100
Db	286	GACTACTTTGAGTCAATGATTTTGCACGATGCAACCAATTAATGATCGAAACGCAAG	345
Qy	101	ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe	120
Db	346	GTCTACTTTGCTGACAAAGATGGTTCCGTTAACTTCCGACCAACCTGTCCAAAGATT	405
Qy	121	LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla	140
Db	406	TGTGAAGCGGACATGTGCGCTTAGACCATATAAAGAAAAACCAATACAAACCAAGCG	465
Qy	141	LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe	160
Db	466	AAATCTGTGTGGAATATATCTGTACAGTTTACTTCCCTTTAAACCTGATGACGATTTC	525
Qy	161	ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr	180
Db	526	AGACCAGGTCTCAAGATATAAGCTATTGAAACACACTAGCTATCGTGACACCATACA	585
Qy	181	SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr	200
Db	586	TCTCAAGAATTAAGTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACG	645
Qy	201	IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu	220
Db	646	ATTATGAACGAGTCTCTCAATCGTCACTCAGCAATGACATTTTCCGTAGCATTTTA	705
Qy	221	ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn	240
Db	706	CCAATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGACCAAGCTTATAGATCAAT	765

Qy	241	LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr	260
Db	766	AAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC	825
Qy	261	ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe	280
Db	826	GTCTTTAAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCATGCACTTGAAACTGTTTC	885
Qy	281	ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr	300
Db	886	ACCATCAATACATGTTGATGTCGATACCAAGAAATGCTAAAAAGTGACGAGCTCTTAACA	945
Qy	301	AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu	320
Db	946	GCTAGCGAAGCTTAACCTTACAGATTTATACGATCTCTGTTAGGCTAAACTA	1005
Qy	321	LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu	340
Db	1006	CTCTACAACAATCTCGATGCTTTTGGTATTATATGACTATACCTTAACCTGGAAGTAGAG	1065
Qy	341	AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly	360
Db	1066	GATAATCAGCATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCCGCAAGGA	1125
Qy	361	GluAsnAlaSerTyrHisLeuAla	368
Db	1126	GAGATGCTAGCTATCATTTAGCC	1149

RESULT 2

US-09-211-542A-5
; Sequence 5, Application US/09211542A
; Patent No. 6210667
; GENERAL INFORMATION:
; APPLICANT: Reed, Guy L.
; TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROMBERG & SUNSTEIN, LLP
; STREET: 125 Summer Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,542A
; FILING DATE: 15-December-1998
; CLASSIFICATION: 1653
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/069,497
; FILING DATE: 15-December-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Attorney, Strimpel, Harriet M.
; REGISTRATION NUMBER: 37,008
; REFERENCE/DOCKET NUMBER: 1874/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 443-9292
; TELEFAX: (617) 443-0004
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1242 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1242
US-09-211-542A-5

Alignment Scores:									
Pred. No.:	2,49e-203	Length:	1242						
Score:	1901.00	Matches:	368						
Percent Similarity:	100.00%	Conservative:	0						
Best Local Similarity:	100.00%	Mismatches:	0						
Query Match:	100.00%	Indels:	0						
DB:	3	Gaps:	0						
US-09-940-235-2_COPY_16_383 (1-368) x US-09-211-542A-5 (1-1242)									
Qy	1	SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu	20						
Db	46	AGCCAAATAGTTGTTAGCGTCTCTGTAATGAGGAGCAATCAAGACATTAGTCTT	105						
Qy	21	LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGly	40						
Db	106	AAATTTTGAATCGATCAACATCACGACCTGCTCATGGAGAAAGACAGAGCAAGC	165						
Qy	41	LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu	60						
Db	166	TAAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTCACATAAACTT	225						
Qy	61	LysAlaAspLeuLeuLysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAsp	80						
Db	226	AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTCATCGCTAAACGTCACAGTAC	285						
Qy	81	AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys	100						
Db	286	GACTACTTTGAGGTCAATGATTTTGAAGCATGCAACCAATTAATCGAAACGGCAAG	345						
Qy	101	ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe	120						
Db	346	GTCTACTTTGCTGACAAAGATGGTTCGGTAACTTGGCGACCCCAACCTGTCCAAAGATTT	405						
Qy	121	LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla	140						
Db	406	TTGCTAAGCGACATGTGCGGCTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG	465						
Qy	141	LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe	160						
Db	466	AAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTT	525						
Qy	161	ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr	180						
Db	526	AGACCAAGTCTCAAGATATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACA	585						
Qy	181	SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr	200						
Db	586	TCTCAAGAAATTAAGTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGGCTATACG	645						
Qy	201	IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu	220						
Db	646	ATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTA	705						
Qy	221	ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn	240						
Db	706	CCAAATGGATCAAGAGTTTACTTACCGTGTGTTAAAAATCGGGAACCAAGCTTATAGGATCAAT	765						
Qy	241	LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr	260						
Db	766	AAAAAATCTGGTCTGAAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC	825						
Qy	261	ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLysLeuPhe	280						
Db	826	GTCTCTTAAAAAAGGGGAAAAACCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTT	885						
Qy	281	ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr	300						
Db	886	ACCATCAAAATACGTTGATGTCATACCAACGAATTCCTAAAAAGTAGCAGCTCTTTAACA	945						
Qy	301	AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu	320						

321	LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu	340	
1006	CTCTACAACAATCTCGATGCTTTTGGTATTATGACTATACCTTAACTGGAAAGATAGAG	1065	
341	AspAsnHisAspAspThrAsnAtqIleIleThrValTyrMetGlyLysArgProGluGly	360	
1066	GATAATCACGATGACACCAACCGTATCATACCGCTTTATATGGCAAGCGACCCGAAGGA	1125	
361	GluAsnAlaSerTyrHisLeuAla	368	
1126	GAGATGCTAGCTATCATTTAGCC	1149	
RESULT 3			
5240845-2			
; Patent No. 5240845			
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO; ; KATANO, TAMIKI;			
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,			
; TSUTOMU			
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS			
; NUMBER OF SEQUENCES: 65			
; CURRENT APPLICATION NUMBER: US/07/549,049			
; FILING DATE: 06-JUL-1990			
; SEQ ID NO.2:			
5240845-2	LENGTH: 1242		
Alignment Scores:			
Pred. No.:	2,49e-203	Length: 1242	
Score:	1901.00	Matches: 368	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	100.00%	Indels: 0	
DB:	6	Gaps: 0	
US-09-940-235-2_COPY_16_383 (1-368) x 5240845-2 (1-1242)			
Qy	1	SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu	20
Db	46	TCCCAGCTGGTGTGTTCCGTAGCTGGCAGCTGTTGAAGTACTAAACAGGACATCTCTCTG	105
Qy	21	LysPheGluIleAspLeuThrSerArgProAlaHisGlyLysThrGluGlnGly	40
Db	106	AAATTTTGAATAATCGACCTCTCGTCCGCCCATGTTGTGTGTAATAACCGAACAGGCG	165
Qy	41	LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu	60
Db	166	CTGTCGCCGAATCTAAACCGTTCGCTACTGACTCTGGCGTATGTCTCATAACTCGAG	225
Qy	61	LysAlaAspLeuLeuLysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAsp	80
Db	226	AAGGCAGATCTGCTGAAAGCAATCCAGGAAACAGCTGATCGCTAACGTTACATTCTTAACG	285
Qy	81	AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys	100
Db	286	GACTACTTTGAGGTAATTCGACTTCGCTAGCGACCTACTATCCCGACCGGTAAACGGCAA	345
Qy	101	ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe	120
Db	346	GTATACTTCGCTGACAAAGACGGTTCGTAACTCTCCGACTCAACCGGTACAGGAATTT	405
Qy	121	LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla	140
Db	406	CTGCTGTCTGGCCATGTACGCGTTTCGCGCGGTACAAAGAAAAACCGATCCAGAACCGGCT	465
Qy	141	LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe	160
Db	466	AAATCTGTTGATGTTGATATACCGTTTCAACCGCTTCAACCGCTGAGATGACTTC	525
Qy	161	ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr	180

Db	946	GCTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTA	1005
Qy	321	LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu	340
Db	1006	CTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATATACCTTAACCTGGAAAGTAGAG	1065
Qy	341	AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly	360
Db	1066	GATAATCACGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCCAGCCGAAGA	1125
Qy	361	GluAsnAlaSerTyrHisLeuAla	368
Db	1126	GAGAAATGCTAGCTATCATTTAGCC	1149
RESULT 3			
5240845-2			
; Patent No. 5240845			
; APPLICANT: FUJII, SETSURO; TAKADA, KAORUKO; KATANO, TAMIKI;			
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,			
; TSUTOMU			
; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS			
; NUMBER OF SEQUENCES: 65			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/07/549,049			
; FILING DATE: 06-JUL-1990			
; SEQ ID NO:2:			
; LENGTH: 1242			
5240845-2			
Alignment Scores:			
Pred. No.:	2,49e-203	Length:	1242
Score:	1901.00	Matches:	368
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-09-940-235-2_COPY_16_383 (1-368) x 5240845-2 (1-1242)			
Qy	1	SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu	20
Db	46	TCCAGCTGGTGTGTTCCGTAGCTGGCACCTGTTGAAGGTACTAACACGAGCATCTCTCTG	105
Qy	21	LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGlnGlnGly	40
Db	106	AAATTTTCAATTCACCTGACCTCTCTGTCGGGCCATCGTGGTAAAAACGACAGGCG	165
Qy	41	LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu	60
Db	166	CTGTCGCCGAAATCTAAACCGTTTCGCTACTGACTCTGGCGCTATGCTCATAAACTCGAG	225
Qy	61	LysAlaAspLeuLysAlaIleGlnGlnGluIleAlaAsnValHisSerAsnAsp	80
Db	226	AAGCGAGATCTGCTGAAAGCAATCCAGGAACAGCTGATCGCTAACGTACATTCTAACGAC	285
Qy	81	AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys	100
Db	286	GACTACTTTGAGGTAAATCGACTTCGCTAGCGACGCTACTATCACCGACCCGTAAACGCAA	345
Qy	101	ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe	120
Db	346	GTATATCTCGTCACAAAGACGGTTCGTAACTCTTCCGACTCAACCGGTACAGGAATTT	405
Qy	121	LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla	140
Db	406	CTGCTGTGGCCATGTACGCGTTCGCCCTCAAAAGAAAAACCGATCCAGAACCCAGGCT	465
Qy	141	LysSerValAspValGluThrValGlnPheThrProLeuAsnProAspAspPhe	160
Db	466	AAATCTGTTGACGTAGAAATACACCGTTCACTTCACTCCCGCTGAAACCCAGACGATGACTTC	525
Qy	161	ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr	180

Db 526 CGCCGGGCTCGAAAGACACTAAACCTGCTGAAAACCTCGGTATCGGTGACACCACTCACT 585
Qy 181 SerGlnGluLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAGGAGCTCCCTGGCTCAGCAGATCTATCTGAAACAAACCATCCGGGCTACACT 645
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 646 ATCTACGAACCGGACTCTTCCATCGTAACCATGACAACGACATCTTCCGTACCACTCTG 705
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 706 CCGATGGACCAAGGAATTTACTTACCGTGTAAACCGCGAACAAGCTTACCGGTATCAAT 765
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 766 AAAAAAATCCGGTCTGAATGAAGAGATTAAACAACACTGACCTGATCTCTGAAAAAGTACTAC 825
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 826 GTACTGAAAAAAGGTGAGAGCGGTATGACCCGTTTCGATCGTTCTCACTGAACTGTTTC 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAAAATACGTTGACGTCGATACCAACGAATTTACTGAACTCTGAGCAGCTGCTGACC 945
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTTCCGAACGTAATCTCGACTTCCGGATCTGTACGACCCCGCTGACAAAGCTAAACATG 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTGTACAAACACCTGGATGCTTTCGGTATCATGACTACACCTGACTGTTGTAAGTAGAA 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GACAACCATGACGACACCAACCGTATCATCCGTTATACATGCGTATACATGGGCAACGTC 1125

RESULT 4
5240845-3
; Patent No. 5240845
; APPLICANT: FUJII, SETSURO; FAKADA, KAORUKO; . KATANO, TAMIKI;
; MAJIMA, EIJI; OGINO, KOICHI; ONO, KENJI; SAKATA, YASUYO; UENOYAMA,
; TSUTOMU

; TITLE OF INVENTION: MUTATED STREPTOKINASE PROTEINS

; NUMBER OF SEQUENCES: 65

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/549,049

; FILING DATE: 06-JUL-1990

; SEQ ID NO:3

; LENGTH: 1262

5240845-3

Alignment Scores:
Pred. No.: 2,55e-203 Length: 1262
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x 5240845-3 (1-1262)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspLysSerLeu 20
Db 60 TCCGAGCTGGTGTGTTCCGTAGCTGGCAGCTGTTGAAGGTACTTAACGAGGACATCTCTCTG 119
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40

Db 120 AAAATTTTCCAAATCGACCTGACCTCTCGTCCGGCCCATGGTGGTAAAAACCGAACAGGGC 179
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValaMetSerHisLysLeuGlu 60
Db 180 CTGTCCCCGAATCTAAACCGTTCCGTACTGACTCTGGGGCTATGTCTCATAACTCGAG 239
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 240 AAGGCAGATCTGCTGAAAGCAATCCAGGAACAGCTGATCGTTAAACGTACATTCTAACGAC 299
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 300 GACTACTTTTGGAGTAATCGACTTCGTAGCAGCGTACTATCACCGACCGTAAACGGCAA 359
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 360 GTATATCTTCGTGACAAAGACCGTTCTGTAACCTCTTCCGACTCAACCGGTACAGGAATTT 419
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 420 CTGCTCTCTGGCCATGTACCGCTTCGCCGTACAAAGAAAAACCGATCCAGAACCGGCT 479
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 480 AAATCTGTTGACGTAGATATACACCGTTCAAGTTCAACCCCGCTGAAACCCAGACGATGACTTC 539
Qy 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 540 CCCCCTGGGCTGAAAGACACTAAACTGTGAAACCCCTGGCTATCCGGTACACCATCACT 599
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 600 TCTCAGGAGCTCTCTGGCTCAGGCACAGTCTATCTCTGAAACAAACCATCCGGGCTACACT 659
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 660 ATCTACGAACCGCATCTTCCATCGTAACCCATGACAACGACATCTTCCGTACCAATCTG 719
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 720 CCGATGGACCGAATTTACTTACCGTGTAAACACCGGAAACGACTTACCCTGATCAAT 779
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 780 AAAAAATCCGGTCTGAATGAAGAGATTAAACAACACTGACTGATCTCTGAAAAAGTACTAC 839
Qy 261 ValLeuLysLysGlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 840 GTACTGAAAAAAGGTGAGAAGCCGTATGACCCGTTTCGATCGTTCTCATCTGAAACTGTTC 899
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 900 ACCATCAAAATAGCTTGGACGTGATACCAACGAATTAATCTGAAAGTCTGAGCAGCTGTCGACC 959
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 960 GCTTCCGAACGTAATCTGGACTTCCCGATCTGTACGACCCCGGTGACAAAGCTAAACTG 1019
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1020 CTGTACAACACCTGGATGCTTTCGGTATCATGGACTACACCCCTGACTGGTAAAGTAGAA 1079
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1080 GACAACCATGACGACACCAACCGTATCATCCCGTATACATGGGCAACGTCGCCGAGGT 1139
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1140 GAAATGCACTCTTACCATCTGGCA 1163

RESULT 5

US-09-211-542A-1

; Sequence 1, Application US/09211542A

Patent No. 6210667
GENERAL INFORMATION:
APPLICANT: Reed, Guy L.
TITLE OF INVENTION: BACTERIAL FIBRIN-DEPENDENT PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROMBERG & SUNSTEIN, LLP
STREET: 125 Summer Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/211,542A
FILING DATE: 15-December-1998
CLASSIFICATION: 1653
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/069,497
FILING DATE: 15-December-1997
ATTORNEY/AGENT INFORMATION:
NAME: Attorney, Strimpel, Harriet M.
REGISTRATION NUMBER: 37,008
REFERENCE/DOCKET NUMBER: 1874/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)443-9292
TELEFAX: (617)443-0004
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2385 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2385
US-09-211-542A-1
Alignment Scores:
Pred. No.: 6.4e-203 Length: 2385
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x US-09-211-542A-1 (1-2385)
Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 1189 AGCCAAATTAGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGAATCAACATTAAGTCCT 1248
Qy 21 LysPheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 1249 AAATTTTGAATCGATTAACATCAGCATCAGCATGCTCATGGAGGAAAGACAGAGCAAGGC 1308
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValMetSerHisLysLeuGlu 60
Db 1309 TTAAGTCAAAATCAAAACCAATTTGCTACTGATGAGGGCGCATGTCACATAAATGAG 1368
Qy 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 1369 AAAGCTGACTTAAGGCTATTCAAGAACCAATTGATCGCTAACCTCCACAGTAACGAC 1428
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 1429 GACTACTTTGAGGTCATTGATTTTGCAGCGATGCAACCACTTACTGATCGAAACGGCAAG 1488

Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 1489 GTCTACTTTTCTGACAAAGATGGTTGCTGTAACCTTGCAGCCCAACCTGTCCAGAAATTT 1548
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 1549 TTGCTAAGCGGACATGTGCGGTTAGACCATATAAGAAACCAATACAAAACCAAGCG 1608
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 1609 AAATCTGTTGATGTGGAATATACTGTACAGTTTACTCTCCCTTAAACCTGATGACGATTC 1668
Qy 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 1669 AGACCAAGTCTCAAGATATAAGCTATTGAAACACATAGCTATCGGTGACACCATCACA 1728
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 1729 TCTCAAGAAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCAACCCAGGCTATACG 1788
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 1789 ATTTATGAAGCTGACTCTCAATCGTCACATCATGACATGACATTTTCCGTACGATTTTA 1848
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaIleArgIleAsn 240
Db 1849 CCAATGGATCAAGAGTTTACTTACCCTGTTTAAATAATCGGAAACAAGCTTATAGGATCAAT 1908
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 1909 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1968
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 1969 GTCTTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGAGTCACTTGAACCTGTTTC 2028
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 2029 ACCATCAATACGTTGATGTCGATACCAACGAATTGCTAAAGAGTGAAGCAGCTCTTAACA 2088
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 2089 GCTAGCGAACGTAACTTAGACTTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTA 2148
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 2149 CTCTACAACAATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACCTGAAAAGTAGAG 2208
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 2209 GATAATCAGATGACACCAACCGTATCATTAACCGTTTATATGGCAAGCGACCCGGAAGGA 2268
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 2269 GAGAATGCTAGCTATCATTTAGCC 2292
RESULT 6
US-08-568-393B-2
Sequence 2, Application US/08568393B
Patent No. 5876999
GENERAL INFORMATION:
APPLICANT: Hua-Lin Wu
APPLICANT: Guey-tueh Shi
TITLE OF INVENTION: Preparation of novel streptokinase
TITLE OF INVENTION: mutants as improved thrombolytic agents
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jeing & Chang
STREET: Two No. 5876999th Second Street, Suite 290
CITY: San Jose
STATE: California
COUNTRY: USA
ZIP: 95113

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1 on Window 3.1
CURRENT APPLICATION DATA:
FILING DATE: US/08/568,393B
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Chi-Ping Chang
REGISTRATION NUMBER: 37,798
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (408) 288-8585
TELEFAX: (408) 288-8386
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1242 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE:
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: SK-K59E
LOCATION: DNA sequence No. 5876999174 and 175 have been changed
LOCATION: from AA to GG, and PROTEIN sequence No. 587699959 has been changed
LOCATION: from Lys to Glu.
OTHER INFORMATION:
US-08-568-393B-2

Alignment Scores:
Pred. No.: 6,996-203 Length: 1242
Score: 1897.00 Matches: 367
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.73% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-08-568-393B-2 (1-1242)

Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 46 AGCCAAATAGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 105
Qy 21 LysPheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 106 AAATTTTGAATCGATCTAATCAATCAGCATCAGCATCGCTCTGAGGAGAAAGACAGAGCAAGGC 165
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 TTAAGTCGGAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTGCACATAAATTGAG 225
Qy 61 LysAlaAspLeuLeuLysAlaLeuGlnGluGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 226 AAAGCTGACTTACTAAAGCTATTCAAGAACAAATTCATCGCTAAACGTCACACGTAACGAC 285
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGTTCATTGATTTTGAACGATGCAACCATTTACTGATCAACGCAAG 345
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 346 GTCTACTTTGCTGACAAAGATGTTCCGTAACCTTCCGCCACCACTGTCCTCAAGATTT 405
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 TTGCTAAGCGGACATGTGCGCTTAGACCATATAAGAAAAACCAATACAAAACCAAGCG 465
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160

Db 466 AAATCTGTTGATGTCGAATATATCTGTACAGATTCTTACTCCCTTAAACCCCTGATCAGATTTC 525
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 AGACCAGGTCTCAAGATACTAAGCTATTGAAACCACTAGCTATCGGTGACACCATCACA 585
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAGAATTACTAGCTCAAGCACAAGCATTTTAAACCAAAACCCACCGCTATACG 645
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 646 ATTTATGAACGTGACTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTA 705
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 706 CCAATGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAAACAAGCTTATAGATCAAT 765
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyrThr 260
Db 766 AAAAAATCTGCTCTGAATGAAGAAATAAACAACACTGACTGATCTCTGGAATAATTATC 825
Qy 261 ValLeuLysGlyGlyLysProTyrAspPropheAspAspSerHisLysLysLeuPhe 280
Db 826 GTCTTTAAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTT 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAAAATAGCTTGTATGCTGATACCAACGAATTCGTAATAAGTGCAGCAGCTCTTAA 945
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTAGCGAACTGAATCTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAACTA 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTCTACAACAATCTCGATGCTTTTGGTATTATGAGTACTATACCTTAACCTGGAAGTAGAG 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GATAATCAGCATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGA 1125
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGATGCTAGTCTATCATTTAGCC 1149

RESULT 7
US-07-854-596B-30
Sequence 30, Application US/07854596B
Patent No. 5434073
GENERAL INFORMATION:
APPLICANT: Dawson, Keith M
APPLICANT: Hunter, Michael G
APPLICANT: Czaplowski, Lloyd G
TITLE OF INVENTION: Proteins and nucleic acids
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dr. John J. McDonnell
STREET: Ten South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/854,596B
FILING DATE: 03-JUN-1992
CLASSIFICATION: 435

100

US-09-940-235-2_COPY_16_383 (1-368) x US-09-374-038-11 (1-1209)

```
QY 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 10 AGCCCAATTAGTTGTTAGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 69
QY 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 70 AAATTTTTCGAATTGACCTTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGC 129
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 130 TTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAA 189
QY 61 LysAlaAspLeuLeuLysAlaLeuGlnGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 190 AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTCATCGCTAACGTCACAGTAACGAC 249
QY 81 AspTyrPheGluValLeuAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 250 GACTACTTTGAGTGCTATTGTTTGCAGCGATGCAACCAATTAATCTGATCGAAACGCAAG 309
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 310 GTCTACTTTGCTGACAAAGATGGTTTCGGTAACCTTGCCGACCCCAACCTGTCGAAGATTT 369
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAla 140
Db 370 TTGCTAAGCGGACATGTGGCGTTAGCATTAAAGAAAAACCAATACAAATCAAGCG 429
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 430 AAATCTGTTGATGTGGAAATATACTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTC 489
```

```
Db 1030 GATAATCACGATGACACCAACCGTATCATAAACGGTTTATATGGGCAAGCGCCGAAGGA 1089
QY 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1090 GAGAAATGCTAGCTATCATTTAGCC 1113
```

RESULT 9

```
US-09-658-179-11
; Sequence 11, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seralena
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-658-179-11
```

```
Alignment Scores:
Pred. No.: 1,95e-200 Length: 1209
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
DB: Gaps: 0
```

US-09-940-235-2_COPY_16_383 (1-368) x US-09-658-179-11 (1-1209)

```
QY 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 10 AGCCCAATTAGTTGTTAGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 69
QY 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 70 AAATTTTTCGAATTGACCTTAACATCAGACCTGCTCATGGAGGAAAGACAGAGCAAGGC 129
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 130 TTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGGATGCCACATAAACTTGAA 189
QY 61 LysAlaAspLeuLeuLysAlaLeuGlnGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 190 AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTCATCGCTAACGTCACAGTAACGAC 249
QY 81 AspTyrPheGluValLeuAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 250 GACTACTTTGAGTGCTATTGTTTGCAGCGATGCAACCAATTAATCTGATCGAAACGCAAG 309
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 310 GTCTACTTTGCTGACAAAGATGGTTTCGGTAACCTTGCCGACCCCAACCTGTCGAAGATTT 369
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAla 140
Db 370 TTGCTAAGCGGACATGTGGCGTTAGCATTAAAGAAAAACCAATACAAATCAAGCG 429
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 430 AAATCTGTTGATGTGGAAATATACTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTTC 489
```


QY 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 490 AGACAGGCTCTCAAGATCTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 549
QY 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 550 TCTCAAGAAATTACTAGTCAAGCAACAAAGCATTTTAAACAAACCCAGCTATACG 609
QY 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 610 ATTTATGAACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTA 669
QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 670 CCAATGGATCAAGAGTTTACTTACCATGTCTCAATCGTCACTCATGACATGACATTTATGAGATCAAT 729
QY 241 LysLysSerGlyLeuAsnGluLeuIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 730 AAAAAATCTGGTCTGAATGAAGAAATAAAACACACTGACCTGATCTCTGAGAAATATTAC 789
QY 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 790 GTCTCTTAAAAAGGGAAAGCGGTATGATCCTTTGATCGCATGCTCACTTGAACCTGTTT 849
QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 850 ACCATCAATACGTTGATGTCAACACCAAGAAATGCTTAAAGACGACGAGCTCTTAACA 909
QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 910 GCTACGCAACGTAACCTAGACTTCAGAGATTTATACGATCCTCGTGTATAGGCTAAACTA 969
QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 970 CTCTACAACTCTCGATGCTTTTGTATTATGAGACTATACCTTAACCTGAAAGTAGAG 1029
QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1030 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 1089
QY 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1090 GAGAATGCTAGCTATCATTTAGCC 1113

RESULT 10

US-07-703-778D-1
; Sequence 1, Application US/07703778D
; Patent No. 5296366
; GENERAL INFORMATION:
; APPLICANT: Garcia, M. P. E. et al
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND EXPRESSION
; TITLE OF INVENTION: OF A GENE WHICH CODES FOR STREPTOKINASE, NUCLEOTIDE SEQUENCE
; TITLE OF INVENTION: OBTAINED, RECOMBINANT DNA AND TRANSFORMED MICROORGANISMS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stanger, Michaelson, Spivak and Wallace, Esq.
; STREET: Parkway 109 Office Center, 328 Newman Springs Road,
; STREET: P. O. Box 8489
; CITY: Red Bank
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2" 1.44Mb IBM compatible diskette
; COMPUTER: IBM PS/2 Model 80
; OPERATING SYSTEM: MS-DOS 5.0
; SOFTWARE: Microsoft Word for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07703,778D
; FILING DATE: 19910522
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Michaelson, Peter L.

; REGISTRATION NUMBER: 30090
; REFERENCE/DOCKET NUMBER: Centro-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)530-6671
; TELEFAX: (908)530-6584
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 base pairs
; TYPE: NUCLEOTIDE WITH CORRESPONDING PROTEIN
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus equisimilis from group C of Lanfield
; ORGANISM: definition
; IMMEDIATE SOURCE: ATCC-9542 strain
; FEATURE: from 1 to 1245 bp mature peptide
; OTHER INFORMATION:
; OTHER INFORMATION: Properties: Streptokinase gene
; OTHER INFORMATION: The gene product binds to human plasminogen
; OTHER INFORMATION: The gene product is an activator of human plasminogen
US-07-703-778D-1
Alignment Scores:
Pred. No.: 2,04e-200 Length: 1245
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x US-07-703-778D-1 (1-1245)
QY 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 46 AGCAATATTAGTTGTAGCGTTGCTGTACTGTGTAGGGGGAGCAATCAAGACATATTAGTCT 105
QY 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 106 AATTTTTTGAATTTGACCTAACATCAGACCTGCTCATGAGGAGAAAGACAGAGCAAGC 165
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 TTAAGTCCAAAATCAAAACCATTTTGTCTACTGTAGTGGCGCATGCCACATAAACTGAA 225
QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnIleAlaAsnValHisSerAsnAsp 80
Db 226 AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGCTCCACGTAACGAC 285
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGGTGCTATGATTTTGAAGCATGCAACCATTTACTGATCGAAGCGCAAG 345
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 346 GTCTACTTTGCTGACAAAGATGTTTCGGTAACCTTGCCGACCCCAACCTGTCCAGAAATTT 405
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 TTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAACCAATACATAAAATCAAGCG 465
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPhe 160
Db 466 AAATCTGTTGATGTGAATATATCTGTACAGTTTACTCCCTTAAACCTCTGATGACGATTTC 525
QY 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 AGACCAAGTCTCAAGATCTAAGTATTGAAACACTAGCTATCGGTGACACCATCACA 585
QY 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAAGAAATTACTAGCTCAAGCACAAAGCATTTTAAACAAACCCACCCTGATACG 645


```
RESULT 12
US-09-658-179-12
; Sequence 12, Application US/09658179
; Patent No. 6413759
; GENERAL INFORMATION:
; APPLICANT: Madrazo, Isis Del Carmen Torrens
; APPLICANT: Garcia, Jose De Jesus De La Fuente
; APPLICANT: Ojalvo, Ariana Garcia
; APPLICANT: Menendez, Alina Seraleña
; APPLICANT: Escalona, Elder Pupo
; APPLICANT: Masso, Julio Raul Fernandez
; APPLICANT: Griego, Martha De Jesus Gonzalez
; TITLE OF INVENTION: STREPTOKINASE MUTANTS
; FILE REFERENCE: Sequence Listings 1-14 re: 976-5
; Patent No. 6413759
; CURRENT APPLICATION NUMBER: US/09/658,179
; CURRENT FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-658-179-12

Alignment Scores:
Pred. No.: 2,04e-200 Length: 1245
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-09-658-179-12 (1-1245)

Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeu 20
Db 10 AGCCAATTAGTTGTTAGCTTCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 69
Qy 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 70 AAAATTTTGAATTTGACCTAATACATCAGCAGCTGCTCATGGAGGAAAGACAGACGAAGC 129
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValMetSerHisLysLeuGlu 60
Db 130 TTAAGTCAAAATCAAAACCACTTCTGCTAGTAGTGGCGCGCATGCCATCAATTAACCTTGA 189
Qy 61 LysAlaAspLeuLeuLysAlaAlaGlnGluGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 190 AAAGCTGACTTAAAGGCTATTCAAGAACAAATTTGATCGCTAACCTCCACAGTAACGAC 249
Qy 81 AspTyrPheGluValLeuAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 250 GACTACTTTGAGGTCAATGATTTTGAAGCGATGCAACCACTTACTGATCGAAACGCAAG 309
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 310 GTCTACTTTGCTGACAAAGATGGTTGGTAACCTTGCCGACCAACCTGTCGAAGAAATTT 369
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 370 TTGCTAAGCGACATGTCGGCGTTAGACCATATTAAGAAACCAATACAAATCAAGCG 429
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 430 AAACTGTTGATGCGAATATCTGATACATTTACTCCCTTAAACCTCATGACGATTTC 489
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 490 AGACCAAGTCTCAAGAGACTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCACA 549
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200

550 TCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCCACCCGCGTATACG 609
201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
610 ATTTATGAAGCTGACTCTCAATCGTCACATCATGACAAATGACATTTTCCGTACGATTTTA 669
221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
670 CCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAAT 729
241 LysLysSerGlyLysLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyr 260
730 AAAAAATCTGGTCTGAATGAAGAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 789
261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
790 GTCTTTAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGAGTCTCTGAAAACCTGTC 849
281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
850 ACCATCAATACGTTGATGTCAACACCAAGAAATTCCTAAAAAGCGAGCAGCTCTTAACA 909
301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
910 GCTAGCGAACGTAACTTAGACTTTCAGAGATTTTATACGATCCTCGCTAAGGCTAAACTA 969
321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
970 CTCTACAACAATCTCGATGCTTTTGGTATTTATGAGACTATACCTTAACTGAAAAGTAGAG 1029
341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
1030 GATAATCAGATGACACCAACCGTATCATAAACGTTTATATGGCAAGCGACCCGGAAGA 1089
361 GluAsnAlaSerTyrHisLeuAla 368
1090 GAGAAATGCTAGCTATCATTTAGCC 1113

RESULT 13
US-07-854-596B-25
; Sequence 25, Application US/07854596B
; Patent No. 5434073
; GENERAL INFORMATION:
; APPLICANT: Dawson, Keith M
; APPLICANT: Hunter, Michael G
; APPLICANT: Czaplowski, Lloyd G
; TITLE OF INVENTION: Proteins and nucleic acids
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. John J. McDonnell
; STREET: Ten South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/854,596B
; FILING DATE: 03-JUN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McDonnell, John J
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 92,337
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
```

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 1257 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..1257

OTHER INFORMATION: /note="Methionyl-streptokinase

OTHER INFORMATION: fusion protein"

FEATURE:

NAME/KEY: CDS

LOCATION: 4..1248

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 4..1248

US-07-854-596B-25

Alignment Scores:

Pred. No.: 2, 07e-200 Length: 1257

Score: 1875.00 Matches: 363

Percent Similarity: 98.91% Conservative: 1

Best Local Similarity: 98.64% Mismatches: 4

Query Match: 98.63% Indels: 0

DB: 1 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-07-854-596B-25 (1-1257)

QY 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20

DB 52 AGCCAAATAGTTGTTAGCGTTGCTGCTGTTGAGGGGACGAATCAAGACATAGTCTT 111

QY 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40

DB 112 AAAATTTTGAATTTGACCTTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGGC 171

QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60

DB 172 TTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGGATGCGACATAAACTTGA 231

QY 61 LysAlaAspLeuLeuValAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80

DB 232 AAAGTGACTTACTAAAGGCTATTCAAGAACAAATGATCGCTAAAGCTCCACAGTACAG 291

QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100

DB 292 GACTACTTTGAGTCAATTTGTTGCAAGCGATGCAACCATTTACTGATCGAAACGCAAG 351

QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120

DB 352 GTCTACTTTGCTGACAAAGATGGTTGCGGTAACTTTGCCGACCAACCTGCTCCAAGAA 411

QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140

DB 412 TTGTTAAGCGACATGTCGGGTTAGACCATATAAAGAAACCAACCAATACAAATCAAG 471

QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160

DB 472 AAATCTGTTGATGGAATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTC 531

QY 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180

DB 532 AGACGAGTCTCAAGATACCTAAGCTATTGAAACACACTAGCTATCGGTGACACCATCA 591

QY 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200

DB 592 TCTCAAGAAATTTACTAGCTCAAGACCAAGACATTTTAAACAAACCAACCATCCAGG 651

QY 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220

DB 652 ATTTATGACGCTGCTCTCAATCTGCTACTCATGACATGACATTTTCCGTAGCATTT 711

QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240

DB 712 CCATGTGATCAAGAGTTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCA 771

QY 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260

DB 772 AAAAAATCTGGTCTGAATGAAGAAATAAACACACCTGACCTGATCTCTCGAATAATTA 831

QY 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280

DB 832 GTCCCTTAAAAAAGGGGAAAGCGGTATGATCCCTTTTGATCGCAGTCACCTTGAAC 891

QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300

DB 892 ACATCAATATAGTTGATGTCAACCAACGAATTTGCTAAAGCGGACGAGCTCTTAACA 951

QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320

DB 952 GCTAGCGAACTTAACCTTACGATTTTATACGATCTCGTGTAAAGGCTAAACTA 1011

QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340

DB 1012 CTCTACACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAGTAGAA 1071

QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360

DB 1072 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCGAAGGA 1131

QY 361 GluAsnAlaSerTyrHisLeuAla 368

DB 1132 GAGATGCTAGCTATCATTTAGCC 1155

RESULT 14

US-07-854-596B-18

; Sequence 18, Application US/07854596B

; Patent No. 5434073

; GENERAL INFORMATION:

; APPLICANT: Dawson, Keith M

; APPLICANT: Hunter, Michael G

; APPLICANT: Czaplowski, Lloyd G

; TITLE OF INVENTION: Proteins and nucleic acids

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. John J. McDonnell

; STREET: Ten South Wacker Drive, Suite 3000

; CITY: Chicago

; STATE: IL

; COUNTRY: USA

; ZIP: 60606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/854,596B

; FILING DATE: 03-JUN-1992

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: McDonnell, John J

; REGISTRATION NUMBER: 26,949

; REFERENCE/DOCKET NUMBER: 92,337

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-715-1000

; TELEFAX: 312-715-1234

; TELEX: 910-221-5317

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1317 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

;
; NAME/KEY: CDS
; LOCATION: 7..1326
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 7..1326
US-07-854-596B-14

Alignment Scores:

Pred. No.: 2,266-200 Length: 1335
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-07-854-596B-14 (1-1335)

Qy	1	SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu	20
Db	130	AGCCAATTAGTTGTTAGGTTGCTGCTGTTGTTAGGGGACGAATCAAGACATTAGTCTT	189
Qy	21	LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly	40
Db	190	AAATTTTTTGAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGGC	249
Qy	41	LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu	60
Db	250	TAAAGTCCAAAATCAAAACCAATTTCTGCTAGTAGTGGCGCATGCCACATAAACTTGA	309
Qy	61	LysAlaAspLeuLeuLysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAsp	80
Db	310	AAAGCTGACTTACNAAGGCTATTCAAGAACAAATTGATCGCTAAAGTCCACAGTAACG	369
Qy	81	AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys	100
Db	370	GACTACTTTGAGGTCAATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCA	429
Qy	101	ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe	120
Db	430	GTCTACTTTGCTGACAAAGATGTTGCGTAACCTTGCCGACCAACCTGTCCAAAGAA	489
Qy	121	LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla	140
Db	490	TTGCTAAGCGACATGTCGGGCTTAGACCATATAAGAAAAACCAATACAAATCAAG	549
Qy	141	LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe	160
Db	550	AAATCTGTTGATGGAATATACTGTACAGTTTACTCCCTTAAACCTGATGACGATTT	609
Qy	161	ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr	180
Db	610	AGACCAGGCTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCA	669
Qy	181	SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr	200
Db	670	TCTCAAGAAATACTAGCTCAAGCAAAAGCATTTTAAACAAACCCATCCAGCTATAC	729
Qy	201	IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu	220
Db	730	ATTTATGAACGTGACTCTCAATCGTCACTCATGACAAATGACATTTCCGTACGATTT	789
Qy	221	ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn	240
Db	790	CCAATGGATCAAGAGTTTACTTACCATGTCAAAATCGGGAACAAGCTTATGAGATCA	849
Qy	241	LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerClulysTyrTyr	260
Db	850	AAAAAATCTGCTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTAC	909
Qy	261	ValLeuLysLysGlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe	280
Db	910	GTCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGGAGTCACCTTGAACCTG	969

Qy	281	ThrIleLysTyrValAspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThr	300
Db	970	ACCATCAATACGTTGATGTTCAACACCAAGATTGCTAAAAAGCGAGCTCTTAACA	1029
Qy	301	AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu	320
Db	1030	GCTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTATAGGCTAAACTA	1089
Qy	321	LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu	340
Db	1090	CTCTACAACAATCTCGATGCTTTTGGTATTATGAGACTATACCTTAACCTGAAAAGTAGAA	1149
Qy	341	AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly	360
Db	1150	GATATACAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGCCCGAAGGA	1209
Qy	361	GluAsnAlaSerTyrHisLeuAla	368
Db	1210	GAGATGCTAGCTATCATTTAGCC	1233

Search completed: November 6, 2004, 04:59:02
Job time : 144.986 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 5, 2004, 21:50:51 ; Search time 5184.77 Seconds
(without alignments)
3356.488 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383
Perfect score: 1901
Sequence: 1 SOLVVSVAGTGTGTQDISL.....IITVYMKRPEGENASYHLA 368

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DB=frames.p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_03112004_174039_11252/app_query.fasta_1.1045
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235_@CGN_1_1_5965_@runat_03112004_174039_11252 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	100.0	1242	6	AR144000
2	1901	100.0	1242	6	AR363845 Sequence
3	1901	100.0	1262	6	E03308 DNA encodin
4	1901	100.0	1262	6	AR363846 Sequence

5	1901	100.0	2385	6	AR143998	AR143998 Sequence
6	1901	100.0	2568	1	STRSKC	K02986 Streptococc
7	1901	100.0	2568	6	A04926	A04926 S.equismil
8	1901	100.0	931	1	SEDEXB	X72832 S.equismil
9	1893	99.6	1401	6	I05204	I05204 Sequence 5
10	1887	99.3	2568	6	E00522	E00522 DNA fragmen
11	1875	98.6	1119	6	I13206	I13206 Sequence 30
12	1875	98.6	1209	6	AR175891	AR175891 Sequence
13	1875	98.6	1209	6	AX030315	AX030315 Sequence
14	1875	98.6	1245	1	S46536	S46536 SKC-2-strep
15	1875	98.6	1245	6	AR175892	AR175892 Sequence
16	1875	98.6	1245	6	AX030316	AX030316 Sequence
17	1875	98.6	1257	6	A20015	A20015 SEQ ID NO:
18	1875	98.6	1257	6	I13203	I13203 Sequence 25
19	1875	98.6	1317	6	A20009	A20009 SEQ ID NO:
20	1875	98.6	1317	6	I13197	I13197 Sequence 18
21	1875	98.6	1335	6	A20006	A20006 SEQ ID NO:
22	1875	98.6	1335	6	I13194	I13194 Sequence 14
23	1875	98.6	1458	6	A20027	A20027 SEQ ID NO:
24	1875	98.6	1458	6	I13215	I13215 Sequence 42
25	1875	98.6	1467	6	A20030	A20030 SEQ ID NO:
26	1875	98.6	1467	6	I13218	I13218 Sequence 46
27	1875	98.6	1512	6	A20016	A20016 SEQ ID NO:
28	1875	98.6	1512	6	I13204	I13204 Sequence 27
29	1875	98.6	2252	6	A20025	A20025 SEQ ID NO:
30	1875	98.6	2253	6	I13213	I13213 Sequence 39
31	1875	98.6	2588	6	A20021	A20021 SEQ ID NO:
32	1875	98.6	2589	6	I13209	I13209 Sequence 34
33	1875	98.6	7057	6	CQ797820	CQ797820 Sequence
34	1871.5	98.4	2566	6	AR068768	AR068768 Sequence
35	1860	97.8	1311	1	AY368335	AY368335 Streptoco
36	1859	97.8	1473	1	SGSKG	X13400 Streptococc
37	1857	97.7	1407	6	E01413	E01413 DNA sequenc
38	1847	97.2	1118	6	A20018	A20018 SEQ ID NO:
39	1819	95.7	1158	6	AR175894	AR175894 Sequence
40	1819	95.7	1158	6	AX030318	AX030318 Sequence
41	1815	95.5	1122	6	AR175893	AR175893 Sequence
42	1815	95.5	1122	6	AX030317	AX030317 Sequence
43	1761	92.6	1323	1	AY234137	AY234137 Streptoco
44	1748	92.0	1323	1	AY234129	AY234129 Streptoco
45	1748	92.0	50354	1	AE014169	AE014169 Streptoco

ALIGNMENTS

RESULT 1	AR144000	Sequence 5 from patent US 6210667.	1242 bp	DNA	linear	PAT 08-AUG-2001
LOCUS	AR144000					
DEFINITION	AR144000					
ACCESSION	AR144000					
VERSION	AR144000.1	GI:15105867				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 1242)					
AUTHORS	Reed,G.L.					
TITLE	Bacterial fibrin-dependent plasminogen activator					
JOURNAL	Patent: US 6210667-A 5 03-APR-2001;					
FEATURES	Location/Qualifiers					
source	1..1242					
	/organism="unknown"					
	/mol_type="unassigned DNA"					
ORIGIN						

Alignment Scores:						
Pred. No.:	1	36e-140	Length:	1242		
Score:	1901.00	Matches:	368			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	100.00%	Indels:	0			
DB:	6	Gaps:	0			

US-09-940-235-2_COPY_16_383 (1-368) x AR144000 (1-1242)

```
QY 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 46 AGCCAAATAGTTGTTAGCGTTGCTGGTACTGTTGAGGCGGCAAGTAACAACATTAGTCTT 105
QY 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 106 AAATTTTGAATCGATTAACATCACGACCTGCTCATGGAGGAAAGACAGACGACGAGGC 165
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 TTAAGTCCAAAATCAAAACCAATTTGCTACTGATGATGGCGCGATGTACATAAATTTGAG 225
QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 226 AAAGCTGACTTAAAGGCTATTCAAGAACCAATTGATCGTTAAACGTCACAGTAAAGC 285
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGTGCTATTGATTTTGCAGCGATGCACCACTTACTGATCGAAACGCGAAG 345
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 346 GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCAGCCCAACCTGTCTCAAGAATTT 405
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 TTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGGC 465
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 466 AAATCTGTTGATGTGAATATATCTGATACGATTTTCCCTTAAACCTGATGACGATTTTC 525
QY 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 AGACAGGCTCTCAAGATACTTAAGCTATTGAAAAACACTAGCTATCGTGCACACCATCACA 585
QY 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAAGAATTTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCCACGAGCTATACG 645
QY 201 IleTyrGluAArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 646 ATTTATGAACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTGACGATTTA 705
QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 706 CCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAAT 765
QY 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 766 AAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 825
QY 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 826 GTCCCTTAAAAAGGGGAAAGCCGTATGATTCCTTTGATCGCAGTCACTTGAACATGTTTC 885
QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAATAGTTGATGTCGATACCAACGAATTTGCTAAAAAGTGAGCAGCTCTTAAACA 945
QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTAGCGAAGCTTAACCTTAGACTTCAGAGATTTTATACGATCTCTGATAAGGCTAAACTA 1005
QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrIleuThrGlyLysValGlu 340
Db 1006 CTCTACACAATCTCGATGCTTTTGGTATATATGGACTATACCTTAACCTGAAAGTAGAG 1065
QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GATATATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAGGA 1125
```

```
QY 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGAAATGCTAGCTATCATTTAGCC 1149
RESULT 2
AR363845 LOCUS 1242 bp DNA linear PAT 03-SEP-2003
DEFINITION Sequence 2 from patent US 5240845.
ACCESSION AR363845
VERSION AR363845.1 GI:34425951
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Fujii,S. deceased, Takada,K. heir, Katano,T., Majima,E., Ogino,K.,
Ono,K., Sakata,Y. and Uenoyama,T.
TITLE Mutated streptokinase proteins
JOURNAL Patent: US 5240845-A 2 31-AUG-1993;
FEATURES Location/Qualifiers
source
1..1242
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores: 1.36e-140 Length: 1242
Pred. No.: 1901.00 Matches: 368
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 0
US-09-940-235-2_COPY_16_383 (1-368) x AR363845 (1-1242)
QY 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 46 TCCAGCTGGTGGTTTCCTGGTAGCTGGCAGCTGTTGAAGGTACTTAACGAGCATCTCTCTG 105
QY 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 106 AAATTTTTCGAATGACCTGACCTCTCGTCCGCCCATCGGTGTAATAACCGAACAGGCG 165
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 CTGTCCCGGAAATCTAAACCGTTTCGCTACTGACTCTGCGCGCTATGTCTATAAATCGAG 225
QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 226 AAGCGAGATCTGCTGAAGCAATCCAGGAACAGCTGATCGTAACTGATCTTAACGAC 285
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGGTAATCGACTTCGCTAGCGAGCTACTATCACCGACCGTAACGCAAA 345
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 346 GTATACTTCGCTGACAAAGACGGTTCTGTAACTCTTCCGACTCAACCGGTACAGGAATTT 405
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 CTGCTCTGTCGGCCATGTACCGTTCCGCCGTACAAAGAAAAACCGATCCAGAACCGAGCT 465
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 466 AAATCTGTTGACCTAGATAACCGTTTCAGTTCACCCCGCTGAACCCAGACGATGACTTC 525
QY 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 CGCCCGGGTCTGAAGACACACTAAACTGCTGAACACCTCGCTATCGGTGACACCATCACT 585
QY 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
```


Db	586	TCTCAGAGAGCTCTGGCTCAGGACAGTCTATCTCTGAACAAAAACCATCCGGGCTACACT	645
Qy	201	IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu	220
Db	646	ATCTACGAACCGCAGCTCTTCATCGTAACCCATGACAAACGACATCTTCCGTACCATCTG	705
Qy	221	ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn	240
Db	706	CCGATGGACCAAGAAATTTACTTACCGTGTAAAAACCGGAAACCAAGCTTACCGTATCAAT	765
Qy	241	LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr	260
Db	766	AAAAAATCCGGTCTGAATGAAGAGATTAAACAACACTGACCTGATCTCTGAAAAGTACTAC	825
Qy	261	ValLeuLysLysGlyGluLysProTyrAspProPheAspAspSerHisLeuLysLeuPhe	280
Db	826	GTACTGAAAANAAGGTGAGAGCCGTATGACCCGTTCCGATCGTTCATCTCATCTGAAACTGTC	885
Qy	281	ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr	300
Db	886	ACCATCAAAATACGTTGACGTCGATACCAACGAATTACTGAAGTCTGAGCAGCTGCTGACC	945
Qy	301	AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu	320
Db	946	GCTTCCGAAACGTAATCTGGACTTCCTCCGCGATCTGTACGACCCGCGTGACAAAGCTTAAACTG	1005
Qy	321	LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu	340
Db	1006	CTGTACAACAACCTGGATGCTTTCGGTATCATGGACTCACCCCTGACTGGTGAAGTAGAA	1065
Qy	341	AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly	360
Db	1066	GACAAACCATGACGACACCAACCGTATCATCACCCTATACATATGCGCAACGTCGCGAAGGT	1125
Qy	361	GluAsnAlaSerTyrHisLeuAla	368
Db	1126	GAATAATGCATCTTACCATCTGGCA	1149
RESULT 3			
LOCUS	E03308	1262 bp	DNA linear PAT 29-SEP-1997
DEFINITION	DNA encoding recombinant streptokinase.		
ACCESSION	E03308		
VERSION	E03308.1	GI:2171525	
KEYWORDS	JP.1992011892-A/1.		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 1262)		
AUTHORS	Fujii,S., Katano,T., Majima,E., Ogino,K., Ono,K., Sakata,Y. and Uenoyama,T.		
TITLE	PROTEIN OF STREPTOKINASES, CORRESPONDING GENE, CORRESPONDING PLASMID RECOMBINANT, CORRESPONDING CHARACTER TRANSFORMANT AND PRODUCTION		
JOURNAL	Patent: JP 1992011892-A 1 16-JAN-1992;		
COMMENT	OTSUKA PHARMACEUT FACTORY INC		
	OS	Artificial gene	
	OC	Artificial sequence; Genes.	
	PN	JP 1992011892-A/1	
	PD	16-JAN-1992	
	PF	06-JUL-1990 JP 1990179851	
	PR	11-JUL-1989 JP 89P 179432, 27-NOV-1989 JP 89P 307957, PR	
	PI	11-APR-1990 JP 90P 96830	
	PI	FUJII SETSUO, KATANO TAMITAKA, MAJIMA EIJI, OGINO KOICHI, PI	
	ONO KENJI,		
	PI	SAKATA YASUYO, UENOYAMA TSUTOMU	
	PC	C12N15/58,C12N1/21,C12N9/70,(C12N1/21,C12R1:19),(C12N9/70, PC	
		C12R1:19);	
	CC	strandedness: Double;	
	CC	topology: Linear;	
	CC	hypothetical: No;	
	CC	anti-sense: No;	

CC	*source: clone=psKX;	Location/Qualifiers	
FH	Key		
FT	5'UTR	1..11	
FT	CDS	12..1256	
FT		/product='recombinant streptokinase'	FT
mat_peptide	12..1253	/product='recombinant streptokinase'	FT
FT	1257..1263.		
3'UTR	Location/Qualifiers		
FEATURES	1..1262		
source	/organism='synthetic construct'		
	/mol_type='genomic DNA'		
	/db_xref='taxon:32630'		
ORIGIN			
Alignment Scores:	1.39e-140	Length: 1262	
Pred. No.:	1901.00	Matches: 368	
Score:	100.00%	Conservative: 0	
Percent Similarity:	100.00%	Mismatches: 0	
Best Local Similarity:	100.00%	Indels: 0	
Query Match:	100.00%	Gaps: 0	
DB:			
US-09-940-235-2_COPY_16_383 (1-368) x E03308 (1-1262)			
Qy	1	SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu	20
Db	60	TCCACGCTGGTGTTCCTGCTAGCTGGCAGCTGTTGAAGGTACTAACCGACATCTCTCTG	119
Qy	21	LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly	40
Db	120	AAATTTTTCGAAATCGACCTGACCTCTCGTCCGCCCATCGTGTAATACCGAACAGGGC	179
Qy	41	LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu	60
Db	180	CTGTCCCGGAAATCTAAACCGTTTCGCTACTGACTCTGGCGCTATGTCTCTATAAATCGAG	239
Qy	61	LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp	80
Db	240	AAGCAGATCTGCTGAAGCAATCCAGGAACACGCTGATCGCTAACGTACATTTCTAAGCAG	239
Qy	81	AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys	100
Db	300	GACTACTTTGAGGTAAATCGACTTCGCTAGCGCGCTACTATCACCGACCGTAAACGCAAA	359
Qy	101	ValTyrPheAlaAspLysAspGlySerValThrIleuProThrGlnProValGlnGluPhe	120
Db	360	GTATACTTCGCTGACAAAGACGGTCTCTGTAACCTCTTCCGACTCAACCGGTACAGGAATTT	419
Qy	121	LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla	140
Db	420	CTGCTGTGTGGCCATGTACGCGTTCGCCGCTACAAAGAAAAACCGATCCAGAACCCAGGCT	479
Qy	141	LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe	160
Db	480	AAATCTGTTGACGTAGATACACCGTTCAGTTCACCCCGCTGAACCCAGACGATGACTTC	539
Qy	161	ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr	180
Db	540	CGCCCGGGTCTGAAGACACTAAACTGCTGAAAACCTCGGTATCGGTGACACCATCACT	599
Qy	181	SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr	200
Db	600	TCTCAGGAGCTCTCGCTCAGGCACAGTCTATCTCTGAACAAAAACCATCCGGGCTACACT	659
Qy	201	IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu	220
Db	660	ATCTAGAACCGGACTCTTCCATCGTAACCCATGACACGACATCTTCCGTACCATCTCTG	719
Qy	221	ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn	240
Db	720	CCGATGGACCAAGAAATTTACTTACCGTGTAAACACCGGCAACAGCTTACCGTATCAAT	779

Qy 241 LysLysSerGlyLeuAsnGluLleAsnAsnThrAspLeuLeuSerGluLysTyrTyr 260
Db 780 AAAAAATCGGTCTGATGAAGAGATTAACAACACTGACTGATCTCTGAAAAGTACTAC 839
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 840 GTACTGAAAAAAGGTGAGAAGCCGTATGACCCGTTTCGATCGTCTCATCTGAAACTGTTC 899
Qy 281 ThrLleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 900 ACCATCAAAATACGTTGACGTCGATACCAACGAATTAACGAAGTCTGAGCAGCTGCTGACC 959
Qy 301 AlaSerGluAtrAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 960 GCTTCGGAACGTAATCTGACATTCGCGATCTGTACGACCCGCGTGAACAAGCTAAACTG 1019
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyLleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1020 CTGTACAAACACTGGATGCTTCGGTATCATGGACTACACCTGACTGGTAAAGTAGAA 1079
Qy 341 AspAsnHisAspAspThrAsnArgLleThrValTyrMetGlyLysArgProGluGly 360
Db 1080 GACAAACCATGACGACACCAACCGTATCATCACCGTATATCATGGCAAAACGTCGGAAGT 1139
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1140 GAAATGCACTTACCATCTGGCA 1163
RESULT 4
LOCUS AR363846 1262 bp DNA linear PAT 03-SEP-2003
DEFINITION Sequence 3 from patent US 5240845.
ACCESSION AR363846
VERSION AR363846.1 GI:34425952
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1262)
AUTHORS Fujii,S. deceased, Takada,K. heir, Katano,T., Majima,E., Ogino,K.,
Ono,K., Sakata,Y. and Uenoyama,T.
TITLE Mutated streptokinase proteins
JOURNAL Patent: US 5240845-A 3 31-AUG-1993;
FEATURES Location/Qualifiers
source 1..1262
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.39e-140 Length: 1262
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x AR363846 (1-1262)
Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspLleSerLeu 20
Db 60 TCCCCAGCTGGTGTTCGCTAGCTGGCACTGTTGAAGGTACTAAACGAGCATCTCTCTG 119
Qy 21 LysPhePheGluLleAspLeuThrSerArgProAlaHisGlyLysThrGluGlnGly 40
Db 120 AAAATTTTTCGAATCGACTGACCTCTGCTCCGGCCCATGGTGTAAACCGAACAGGC 179
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 180 CTGTCCCCGAAATCTAAACCGTTCGCTACTGACTCTGGCGCTATGTCTCATAAATCGAG 239
Qy 61 LysAlaAspLeuLeuLysAlaLleGlnGluGlnLeuLleAlaAsnValHisSerAsnAsp 80

Db 240 AAGGCAGATCTGCTGAAAGCAATCCAGGAACAGCTGATCGCTAAACGTACATTTCTAACGCAC 299
Qy 81 AspTyrPheGluValLleAspPheAlaSerAspAlaThrLleThrAspArgAsnGlyLys 100
Db 300 GACTACTTTTGGAGTAATCGACTTCGCTAGCGACGCTACTATCACCCGACCGTAAACGGCAA 359
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 360 GTATACTTTCGCTGACAAAGACGGTCTCTGTAACCTCTTCGACTCAACCGGTACAGGAATTT 419
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProLleGlnAsnGlnAla 140
Db 420 CTGCTCTCTGGCCATGTACGCGTTCGCCGCTCAAAAGAAAAACCGATCCAGAACCAAGCT 479
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPhe 160
Db 480 AAATCTGTTCGCTGACGTAGAATACACCGTTTCAGTTTCCACCCGCTGAACCCGACGATGACTTC 539
Qy 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaLleGlyAspThrLleThr 180
Db 540 CGCCCGGGTCTGAAAGACACTAAACTGCTGAAACCCCTGGCTATCGGTGACACCATCACT 599
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerLleLeuAsnLysAsnHisProGlyTyrThr 200
Db 600 TCTCAGGAGCTCCTGGCTCAGGCACAGTCTATCTCTGAACAAAAACCATCCGGCTACACT 659
Qy 201 IleTyrGluArgAspSerSerLleValThrHisAspAsnAspLlePheArgThrLleLeu 220
Db 660 ATCTACGAACCGGACTCTTCCATCGTAACCCATGACACGACATCTTCCGTACCATCTCTG 719
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgLleAsn 240
Db 720 CCGATGGACCGAGAAATTTACTTACCGTGTGTAAACCCGCAACAGCTTACCGTATCAAT 779
Qy 241 LysLysSerGlyLeuAsnGluLleAsnAsnThrAspLeuLleSerGluLysTyrTyr 260
Db 780 AAAAAATCCGGTCTGAATGAAGAGATTAACAACACTGACTGATCTCTGAAAAGTACTAC 839
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 840 GTACTGAAAAAAGGTGAGAAGCCGTATGACCCGTTTCGATCGTCTCATCTGAAACTGTTC 899
Qy 281 ThrLleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 900 ACCATCAAAATACGTTGACGTCGATACCAACGAATTAACGAAGTCTGAGCAGCTGCTGACC 959
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 960 GCTTCCGAAACGTAATCTGGACTTCGCGATCTGTACGACCCGCGTGAACAAGCTAAACTG 1019
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyLleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1020 CTGTACAAACACTGGATGCTTTTCGCTATCATGGACTACACCCCTGACTGTTGTAAGTAGAA 1079
Qy 341 AspAsnHisAspAspThrAsnArgLleLleThrValTyrMetGlyLysArgProGluGly 360
Db 1080 GACAAACCATGACGACACCAACCGTATCATCACCGTATATCATGGGCAAAACGTCGGAAGT 1139
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1140 GAAATGCACTTACCATCTGGCA 1163
RESULT 5
LOCUS AR143998 2385 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6210667.
ACCESSION AR143998
VERSION AR143998.1 GI:15105865
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 2385)
AUTHORS Reed,G.L.
TITLE Bacterial fibrin-dependent plasminogen activator
JOURNAL Patent: US 6210667-A 1 03-APR-2001;
FEATURES Location/Qualifiers
source 1..2385
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2.8e-140 Length: 2385
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AR143998 (1-2385)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 1189 AGCCAATTAGTTGTTAGCGTTGCTGCTACTGTTGAGGGACGAATCAAGACATTAGTCTT 1248
Qy 21 LysPheGluLeuLeuAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 1249 AAAATTTTGAATCGATCTAACATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGGC 1308
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 1309 TTAAGTCCTCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAATTTGAG 1368
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuLeuIleAlaAsnValHisSerAsnAsp 80
Db 1369 AAAGCTGACTTACTAAGAGCTATTCAAGAACAAATGATCGTAACTCCACAGTAACGAC 1428
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 1429 GACTACTTTGAGTTCATGATTTTGCACGGATGCAACCATTAATGATGCAACGCAAG 1488
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 1489 GTCTACTTTGCTGACAAAGATGGTTGGGTAACTTGCACCACTTGCACCACTGCTCCAAAGATTT 1548
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAla 140
Db 1549 TTGCTAAGCGGACATGTGCGCGGTAGACCATATAAGAAAAACCAATACAAACCAAGCG 1608
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 1609 AAATCTGTTGATGCGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTC 1668
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 1669 AGACCAGGCTCTCAAGACTAAGCTATTGAAACACTAGCTATCGTGACACCATCACCA 1728
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 1729 TCTCAAGAATTTACTAGCTCAAGCACCAAGCAATTTTAAACAAACCAACCCAGGCTATACG 1788
Qy 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 1789 ATTTATGAACGTGATCTCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTA 1848
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 1849 CCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACCAAGCTTATAGATCAAT 1908
Qy 241 LysLysSerGlyLeuAsnGluGluLeuAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 1909 AAAAAATCTGGTCTCAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1968
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280

Db 1969 GTCCTTAAAAAAGGGAAAAAGCCGATGATGATCCCTTTGATCGACTCACTTGAACACTGTTTC 2028
Qy 281 ThrIleLysTyrValAspValAspThrAsnGlnLeuLeuLysSerGluGlnLeuLeuThr 300
Db 2029 ACCATCAATACGTTGATGTCGATACCAACGAATGCTAAANAAGTGACAGCTCTTAAACA 2088
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 2089 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTATAGCGTAAACTA 2148
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 2149 CTCTACAAACATCTCGATGCTTTTGGTATTATGGAATATACCTTAACCTGGAAGAGTAGAG 2208
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 2209 GATATATCAGATGACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 2268
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 2269 GAGAATGCTAGCTATCATTTTAGCC 2292
RESULT 6
STRSKC
LOCUS Streptococcus equisimilis (H46A) DNA linear BCT 26-APR-1993
DEFINITION Streptococcus equisimilis (H46A) streptokinase gene, complete cds.
ACCESSION K02986
VERSION K02986.1 GI:153808
KEYWORDS streptokinase.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus dysgalactiae subsp. equisimilis
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 2568)
Malke,H., Roe,B. and Ferretti,J.J.
Nucleotide sequence of the streptokinase gene from Streptococcus
equisimilis H46A
JOURNAL Gene 34 (2-3), 357-362 (1985)
MEDLINE 85232082
PUBMED 2989113
COMMENT Original source text: S.equisimilis (strain H46A) DNA, clone pMF5.
Draft entry and hard copy of sequence for [1] kindly provided by
J.J.Ferretti, 03-SEP-1985.
The -35 and -10 regions are located at positions 760-765 and
781-786 respectively and an SD sequence at 808-813. Downstream
from the coding region inverted repeats (positions 2176-2190 and
2203-2217) are thought to function as transcription terminators.
The nucleotide sequence of skc does not support the hypothesis that
the gene has evolved by duplication and fusion, as suggested by
internal two-fold AA homologies of its product.
FEATURES
Location/Qualifiers
source 1..2568
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="genomic DNA"
/db_xref="taxon:119602"
mRNA 794..22141
/product="skc mRNA"
CDS 819..2141
/notes="prestreptokinase"
/codon_start=1
/transl_table=11
/protein_id="AAA26974.1"
/db_xref="GI:153809"
sig_peptide 819..896

```
/note="streptokinase signal peptide"
mat_peptide      897..2138
                  /product="streptokinase"
ORIGIN           5 bp upstream of PstI site.

Alignment Scores:
Pred. No.:      3 04e-140      Length:      2568
Score:          1901.00        Matches:     368
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     100.00%      Indels:       0
DB:              1           Gaps:         0

US-09-940-235-2_COPY_16_383 (1-368) x STRSKC (1-2568)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 942 AGCCAAATTAGTTGTTAGCGTTCTGCTACTGTTGAGGGACGAATCAAGACATTAAGTCTT 1001

Qy 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 1002 AAATTTTTTGAATCGAATCAATCAGACCTGCTCATGGAGGAAAAGACAGAGCAAGGC 1061

Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 1062 TTAAGTCCAAATCAAAACCAATTGCTACTGATAGTGGCGCGATGTCACATAAAGTTGAG 1121

Qy 61 LysAlaAspLeuLysAlaLeuGlnGluLeuLeuAlaHisValMetSerHisLysLeuGlu 80
Db 1122 AAAGCTGACTTACTAAAGGCTATTCAAGAACCAATTGATCGCTAAACGTCACAGTAACGAC 1181

Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgHisGlyLys 100
Db 1182 GACTACTTTGAGGTCAATGATTTTGAAGCGATGCAACCAATTACTGATGAAACGGCAAG 1241

Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 1242 GTCTACTTTGCTGACAAAGATGTTGCGTAACTTGGCCGCCACCACTGTCCTCAAGAAATTT 1301

Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 1302 TTGCTAAGCGGACATGTGCGCTTAGACCATATAAGAAAAACCAATACAAAAACCAAGCG 1361

Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 1362 AAATCTGTTGATGTGGAATATACTGTACAGTTTACTTCCCTTAAACCCCTGATGACGATTC 1421

Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 1422 AGACCAGGTCTCAAGATACTTAAGCTATTGAAACACTAGCTATCGTGACCATCACACA 1481

Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 1482 TCTCAAGAAATTACTAGCTCAAGCAACAAAGCAATTTAAACAAAAACCAACCCAGGCTATACG 1541

Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 1542 ATTTATGAACGTCGACTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTA 1601

Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 1602 CCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAAT 1661

Qy 241 LysLysSerGlyLeuAsnGluGluLeuAsnThrAspLeuLeuSerGluLysTyrTyr 260
Db 1662 AAAAAATCTGGTCTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTAC 1721

Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLysLeuPhe 280
Db 1722 GTCCCTTAAAAAAGGGGAAAGCCGATATGATCCCTTTGATCGCAGTCACCTTGAACATGTTTC 1781

Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 1781
```

Db 1242 GTCTACTTTGTCACAAAGATGGTTCGTAACCTTCCGACCAACCTGTCTCCAAAGATTT 1301
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAenGlnAla 140
Db 1302 TTGCTAAGCGGACATGTCGGCTTAGACCATATAAGAAAACCAATACAAAACCAAGCG 1361
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 1362 AAATCTGTTGATGCGAATATACTGTACAGTTTACTCCCTTAAACCTGATGAGATTTC 1421
Qy 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 1422 AGACAGAGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 1481
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 1482 TCTCAAGAAATTAAGTCAAGCACAAAGCATTTTAAACAAAACCCAGCGCTATACG 1541
Qy 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 1542 ATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTA 1601
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 1602 CCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAAT 1661
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyrTyr 260
Db 1662 AAAAATCTGGTCTGAATGAAGAAATAAACACACCTGACCTGATCTCTGAGAAATATTAC 1721
Qy 261 ValLeuLysLysGlyLysProTyrAspProPheAspAspArgSerHisLeuLysLeuPhe 280
Db 1722 GTCCCTTAAAAAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTT 1781
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 1782 ACCATCAATACGTTGATGTCGATACCAAGCAATGCTTAAAGAGTCAGCAGCTCTTAAACA 1841
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 1842 GCTAGCAACGTAACCTTAGACTTCAGAGATTTTATACGATCTCTCGTAAAGGCTAAACTA 1901
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1902 CTCTACAAATCTCATGCTTTTGGTATTATGGACTATACCTTAACTGAAAGTAGAG 1961
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1962 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 2021
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 2022 GAGAATGCTAGCTATCATTTAGCC 2045
RESULT 8
SEDEXB
LOCUS SEDXB 8931 bp DNA linear BCT 17-FEB-1997
DEFINITION S.equisimilis dextb, abc, lrp, skc, rel genes and ORF1.
ACCESSION X72832
VERSION X72832.1 GI:407876
KEYWORDS abc gene; ABC transporter; dextB gene; dextran glucosidase; glucan 1,6-alpha-glucosidase; leucine rich protein; LRP gene; ORF1; rel gene; skc gene; streptokinase; stringent response-like protein.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Streptococcus
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 3621 to 6190)
AUTHORS Malke,H., Roe,B. and Ferretti,J.J.
TITLE Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H46A
JOURNAL Gene 34 (2-3), 357-362 (1985)

MEDLINE 85232082
PUBMED 2989113
REFERENCE 2 (bases 1 to 4188; 5790 to 8931)
AUTHORS Mechold,U., Steiner,K., Vettermann,S. and Malke,H.
TITLE Genetic organization of the streptokinase region of the Streptococcus equisimilis H46A chromosome
Mol. Gen. Genet. 241 (1-2), 129-140 (1993)
MEDLINE 94049672
PUBMED 8232196
REFERENCE 3 (bases 1 to 8931)
AUTHORS Malke,H.
TITLE Direct Submission
SUBMITTED (05-MAY-1993) H. Malke, Institute for Molecular Biology, Jena University, Winzerlaer Str 10, 07708 Jena, FRG
COMMENT Related sequences: K02986, M19346, X13399 & X13400.
FEATURES
Location/Qualifiers
1..8931 /organism="Streptococcus dysgalactiae subsp. equisimilis"
2 (bases 1 to 4188; 5790 to 8931) /mol_type="genomic DNA"
/strain="H46A"
/isolates="human group C strain"
/sub_species="equisimilis"
/db_xref="taxon:119602"
/chromosome="streptokinase region"
/clone_lib="E.coli plasmid library containing subfragments of the submitted seq: pSHD14/16, pSPV19, pSH2, pME1, pCWP73, pRH10, pWX4"
gene complement(89..1761)
terminator /gene="dextB"
complement(89..115)
/gene="dextB"
/note="hairpin loop"
complement(136..1749)
CDS /gene="dextB"
/EC_number="3.2.1.70"
/codon_start=1
/transl_table=11
/product="glucan 1,6-alpha-glucosidase"
/protein_id="CAA51348.1"
/db_xref="GI:407877"
/db_xref="GOA:Q59905"
/translation="MQQWKKATIIQIYPRSKTSGNGIGLKGITSLDYLQKLG ITAIIWSPVQSPMDNGYDISDYEAIAEVFGNMDMDLLAAANERGIKIIMDLVNV HTSDEHAMFVEARENPNSPERDVIWRDENNLWSIFSGAWELDSAGQYIHLFSK KQDLNWEAHVRQKIYDMNFWIAKIGIGCFMDVIDLIGKIPDSITNGPRLHDIY KENQATFGNHDVMTVGTEATGATFARQYSRPNKELSMFOFEHVGLQHPNAPKW DYAEELDVPAKTIIFSKQTELKLGWNSLFWNHDLPVLVSIWGNDSIYRKSASKA LATLLHMRGTPYIQGEIGITMYPFKDLTEVDIESLNYAKAEMENGYPARVMSS IRKVGDRNARTPMQSKDTHAGFSEAQETLVPVNPVQEIIVADALANODSIPTVYQQ LIALRKQDWLVEADYHLLPTADKVFAYQRFQGEETVYIVVNVSDQEQVPAKDLAGE VVITNTDVKVLETRHLPQWDAFCVKLSV"
RBS complement(1757..1761)
gene /gene="dextB"
complement(1780..3051)
terminator /gene="abc"
complement(1780..1810)
/gene="abc"
/note="hairpin loop"
complement(2482..2499)
misc_feature /gene="abc"
/note="Walker motif B"
complement(2633..2965)
CDS /gene="abc"
/codon_start=1
/transl_table=11
/product="ABC transporter"
/protein_id="CAA51349.1"
/db_xref="GI:600072"
/db_xref="GOA:Q54086"
/db_xref="TrEMBL:Q54086"
/translation="MVELNLNHIYKKYPNTHYVAVDFOLDIKDKKEFIYVFGSPGGCK STTLRMIAGLEIDISEGELKIDGEVNDKSPKDRDIAMVFQNYALYPHMTYVDNMAFGL

```

misc_feature
/feature_id="KLRKYKXR"
/complement(2644)
/gene="abc"
/note="(+)1 frame shift mutation in H46A"
misc_feature
/complement(2828..2854)
/gene="abc"
/note="Walker motif A"
RBS
/complement(2973..2976)
/gene="abc"
-10_signal
/complement(3026..3031)
/gene="abc"
-10_signal
/complement(3043..3048)
/gene="abc"
misc_feature
/complement(3050..3051)
/gene="abc"
/note="TG motif"
gene
/complement(3244..4149)
/gene="lrp"
terminator
/complement(3244..3261)
/gene="lrp"
CDS
/note="hairpin loop"
/complement(3262..4110)
/gene="lrp"
/codon_start=1
/transl_table=11
/product="leucine rich protein"
protein_id="CAA51350.1"
/db_xref="GI:407878"
/db_xref="GOA:Q54087"
/db_xref="Swiss-Prot:O54087"
translation="MELKDYFEMQVGHPLGDKWVSEKGVYVHFKSCLSKER
LLEVLGQYEVQLQSGSPWQRILLDHQNPQLPETSQFIYLNHOQVLPADLVLLQ
VMIAGLEVILPISITQTAPLCROATSIKVLRSLEGLLPLESDFGALTFMVGNAWYQ
QAAGTIRCFEECOLLTAYLKQSGGKLLTFAEVLWSILSHQSPFALTROFHPLN
POSMDVVVHALWSEHNLVQTAQRLYIHNSLQYKLDKFAQOQSLHLKQLDLAFAY
LFLKKY"
misc_feature
/complement(3337..3423)
/gene="lrp"
/note="leucine zipper motif"
RBS
/complement(4117..4120)
/gene="lrp"
-10_signal
/complement(4141..4146)
/gene="lrp"
misc_feature
/complement(4148..4149)
/gene="lrp"
/note="TG motif"
misc_feature
/4172..4368
/note="static DNA bending locus"
gene
/4392..5837
/gene="skc"
misc_feature
/4392..4393
/gene="skc"
/note="TG motif"
-10_signal
/4395..4400
/gene="skc"
-10_signal
/4401..4406
/gene="skc"
/note="alternative"
RBS
/4428..4431
/gene="skc"
CDS
/4439..5761
/gene="skc"
/codon_start=1
/transl_table=11
/product="streptokinase"
protein_id="CAA51351.1"
/db_xref="GI:407879"
/db_xref="GOA:P00779"
/db_xref="Swiss-Prot:P00779"
translation="WKNYLSFGMFAALLFALTGTNTVNSVOAIAGPEWLLDRPNSVNSQL
VVSVAAGTVEGTNQDLSLKFEFDLTSRPAHGKTEQGLSPKSPKPAATDSGAMSHKLEK
ADLLKAIQEQLIANVHSNDDYFEVIDPASDATITDRNGKVPADKDGSVTLPTQPVQE
FLLSGHRVVRPYKEKPIQNAKSDVVEYTVQFTPLNPDDDFRGLKDTLLKTLAIGD
TTTSOELLAAQOSIILNKNHPGYTIYERDSIVTHDNDIFRILPMPQDEFTYRKNREQ
AVRINKSGLINEINNTDLISEKYVYLKKEGYPDPDFDRSHLKLFTIKYVDVDTNELL
KSEQLLTASERNLDPRDLYDRPKAKLLYNNLDAFGIMDYTLTGKVEDNHDHNRILT
VWGKRPEGENASVHLAYDKDRYTEEREVYSYLYRTGTPTIPDPNPK"
sig_peptide
/feature_id="sig"
terminator
/5796..5837
/gene="skc"
CDS
/complement(5859..6302)
/note="unnamed protein product; ORF1"
/codon_start=1
/transl_table=11
/protein_id="CAA51352.1"
/db_xref="GI:407880"
/db_xref="GOA:Q54088"
/db_xref="Swiss-Prot:O54088"
translation="MKLVLRQKEASVSDGKIAGAINOGLLLLVGVGPDPAEDLAY
AVRKIVNMRIFSDADGKMNSIQDIKGSILSVSQFTLYADTKGKGNRPAPFTGAAPDMA
SOFYDRFNEQLADFVPVERGVFGADMQVSLINDGPVTIILDTKCH"
conflict
/6007..6009
/citation=[1]
/replacement="gc"
6162..6164
/citation=[1]
/replacement="tg"
RBS
/complement(6310..6313)
/note="ORF1"
gene
/complement(6317..8575)
/gene="rel"
CDS
/complement(6317..8536)
/gene="rel"
/codon_start=1
/transl_table=11
/product="stringent response-like protein"
protein_id="CAA51353.1"
/db_xref="GI:407881"
/db_xref="GOA:Q54089"
Alignment Scores:
Pred. No.: 1.2e-139 Length: 8931
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x SEDEXB (1-8931)
QY 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 4562 AGCCAAATTAGTTGTAGCGTTGCTGTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 4621
QY 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 4622 AAATTTTGGAAATCGATCTTAACATCAGACCTGCTCATGGAGAGAAAGACAGAGCAGGC 4681
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 4682 TTAAGTCCAAATCAAAACCATTTGCTACTGATGTGGCGGATGTCCACATAAATTGAG 4741
QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaValHisSerAsnAsp 80
Db 4742 AAAGCTGACTTAAAGGCTATTCAAGAACAAATTGATCGTAAACGTCACATAACGAC 4801
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 4802 GACTACTTTGAGTCAATTGATTTTTCAGCGATGCAACCACTTACTGATCGAAACGCAAG 4861
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 4862 GTCTACTTTGTCGACAAAGATGCTTCGTAACCTTCCGACCCCAACCTGCTCCAGAAATTT 4921

```

```
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 4922 TTGCTTAAGCGGACATGTGCGGTAGACCATATAAAGAAACCAATACAAAACCAAGCG 4981
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 4982 AAATCTGTTGATGTGAATATACTGTACAGTTTACTCCCTTAAACCTGATGAGATTTC 5041
QY 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 5042 AGACACAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 5101
QY 181 SerGlnGluLeuLeuAlaGlnGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 5102 TCTCAAGATTACTAGCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATACG 5161
QY 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 5162 ATTTATGAACGTGACTCCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTA 5221
QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 5222 CCAATGGATCAAGAGTTTACTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAAT 5281
QY 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 5282 AAAAAATCTGGTCTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAAATATTAC 5341
QY 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 5342 GTCCCTTAAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAAACTGTT 5401
QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 5402 ACCATCAATATAGTTGATGTGCATACCAACGAATTCGCTAAAGAGTGAGCGCTCTTAACA 5461
QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 5462 GCTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTA 5521
QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 5522 CTCTCAACAATCTCGATGCTTTTGGTATTATGGAATACTATACCTTAACCTGGAAGTAGAG 5581
QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 5582 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGA 5641
QY 361 GluAsnAlaSerTyrHisLeuAla 368
Db 5642 GAGATGCTAGCTATCACTTTAGCC 5665
```

RESULT 9

```
105204
LOCUS I05204 1401 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 5 from Patent EP 0248227.
ACCESSION I05204
VERSION I05204.1 GI:591209
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
```

REFERENCE

```
1 (bases 1 to 1401)
AUTHORS Hagenson,M.J. and Stroman,D.W.
TITLE Yeast production of streptokinase
JOURNAL Patent: EP 0248227-A1 5 09-DEC-1987;
FEATURES Location/Qualifiers
```

source

```
1..1401
/organism="unknown"
/mol_type="unassigned DNA"
```

ORIGIN

Alignment Scores:

```
Pred. No.: 6,63e-140 Length: 1401
Score: 1893.00 Matches: 366
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 2
Query Match: 99.58% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x I05204 (1-1401)

QY 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 55 AGCNAATTAGTTGTTAGCGTTGCTGTTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 114
QY 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 115 AAATTTTTCGAAATCGATCTTAACATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGC 174
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 175 TTAAGTCCAAAATCAAAACCAATTTTGTCTACTGATAGTGGCGCGATGTCCACATAAATTTGAG 234
QY 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnIleAlaAsnValHisSerAsnAsp 80
Db 235 AAAGCTGACTTACTTAAAGGCTATTTCAGAAACAATTGATCGCTTAACGTCACAGTAAGCAC 294
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 295 GACTACTTTGAGGTCAATGATTTTGCACGCGATGCAACCACTTACTGATCGAAGCGCAAG 354
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 355 GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCCAACCTGTGCCAAGAATTT 414
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 415 TTGCTTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAACCAATACAAAACCAAGCG 474
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 475 AAATCTGTTGATGTGAATATATCTGTACAGTTTACTCCCTTAAACCTGATGACGATTTC 534
QY 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrIleuAlaIleGlyAspThrIleThr 180
Db 535 AGACCAAGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 594
QY 181 SerGlnGluLeuLeuAlaGlnGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 595 TCTCAAGAATTACTAGCTCAAGCACAAAGCAATTTTAAACAAACCAACCCAGGCTATACG 654
QY 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 655 ATTTATGAACGTGACTCCCTCAATCGTCATCATGACAATGACATTTTCCGTACGATTTTA 714
QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 715 CCAATGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAAT 774
QY 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 775 AAAAAATCTGGTCTGAATGAAGAAATAAACACACTGACCTGATCTCTGAGAATAATTAC 834
QY 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 835 GTCCCTTAAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACCTGTTTC 894
QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 895 ACCATCAATATAGCTTGTGTCGATACCAAGCAATTCCTAAAGAGTGAGCGAGCTCTTAACA 954
QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 955 GCTAGCGAAGCTAACTTAGACTTCAGAGATTTTATACGATCTCTCGTGATAAGGTCAAACTA 1014
```

```
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
D 1015 CTCTACAAACATCTCGATGCTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1074
Qy 341 AspAsnHisAspAspThrAsnArgIlelleThrValTyrMetGlyLysArgProGluGly 360
D 1075 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 1134
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
D 1135 GAGAAATGTCAGCTATCAATTAGCC 1158

RESULT 10
E00522
LOCUS E00522 2568 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA fragment of plasmid PMF1 into which DNA encoding streptokinase
is inserted.
ACCESSION E00522
VERSION E00522.1 GI:2168801
KEYWORDS JP 1985237995-A/1.
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 2568)
AUTHORS Jiyosefu J.F. and Horusuto, M.
TITLE RECOMBINATION VECTOR FOR PRODUCING STREPTOKINASE
JOURNAL Patent: JP 1985237995-A 1 26-NOV-1985;
PHILLIPS PETROLEUM CO
OS Streptococcus equisimilis
PN JP 1985237995-A/1
PD 26-NOV-1985
PF 09-OCT-1984 JP 1984212403
PR 10-OCT-1983 DD 83 255523, 02-MAR-1984 US 84 585417 PI
JIXOSEFU JIEI FUERTSUTSI, HORUSUTO MARUKE
PC C12N15/00, C12N1/20, C12N9/70, (C12N15/00, C12R1:46), (C12N1/20, PC
C12R1:19),
PC (C12N9/70, C12R1:19);
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=Streptococcus equisimilis H46A; CC *source:
library=streptococcus equisimilis H46 library; CC *source:
clone=lamda L47 skc clone;
FH Key Location/Qualifiers
FT 5'UTR 1..818
FT sig_peptide 819..896
FT /product='streptokinase signal peptide' FT
mat_peptide 897..2138
FT /product='streptokinase'
FT CDS 819..2143
FT /product='streptokinase'
FT 3'UTR 2142..2569.
FEATURES
source
1..2568
/organism='Streptococcus dysgalactiae subsp. equisimilis'
/mol_type='genomic DNA'
/db_xref='taxon:119602'

ORIGIN
Alignment Scores:
Pred. No.: 3,84e-139 Length: 2568
Score: 1887.00 Matches: 366
Percent Similarity: 99.46% Conservative: 0
Best Local Similarity: 99.46% Mismatches: 2
Query Match: 99.28% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x E00522 (1-2568)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
```

```
Db 942 AGCCAATTAGTTGTTAGCGTTGCTGTACTGTTAGGGGACGAATCAAGACATTAGTCTT 1001
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyIleThrGluGlnGly 40
D 1002 AAATTTTTTGAATTCGATCTAAACATCAGCCTGCTCATCGGAGAAAGACAGACGAAGC 1061
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
D 1062 TTAAGTCCAAATCAAAACCATTTGCTACTGATGTGCGCGATGTGCATATAACTTGAG 1121
Qy 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
D 1122 AAAGCTGACTTAAAGGCTATTCAAGAACAAATTCATCGCTAAACGCTCCACAGTAACGAC 1181
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
D 1182 GACTACTTTTGAGTCTATTGATTTGCAAGCGATGCAACCATTTACTGATCGAAACGCCAAG 1241
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
D 1242 GTCCTACTTTGCTGACAAAGATGTTCCGTAACCTTGCCGACCCCAACCTGTCGAAGATTT 1301
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
D 1302 TTGCTAAGCCACATGTCGCGGTAGACCATATAAGAAAAACCAATACAAAAACCAAGCG 1361
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPhe 160
D 1362 AAATCTGTTGATGTGAATATATCTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTC 1421
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
D 1422 AGACCAAGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCCGTGACACCATACACA 1481
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
D 1482 TCTCAAGATTAAGTCTCAAGCACACACAGCAATTTTAAACAAAAACCAACGAGCTATACG 1541
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
D 1542 ATTATGAACGTCAGCTCCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTA 1601
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
D 1602 CCAATGGATCAAGAGTTTACTTACCGTGTATAAAATCGGGAAACAGCTTATAGGATCAAT 1661
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
D 1662 AAAAATCTGGTCTGAATGAAGAAATAAACACACTGACTGATCTCTGAGAAATATTAC 1721
Qy 261 ValLeuLysGlyGlyLysProTyrAspProPheAspAspArgSerHisLeuLysLeuPhe 280
D 1722 GTCCCTTAAAAAGGGGAAAGCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTC 1781
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
D 1782 ACCATCAAAATAGCTGTGATACCAACGAATTCGTAAGAAAGTGAAGAGCTCTTTAAACA 1841
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
D 1842 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAACTA 1901
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
D 1902 CTCTACAAACATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1961
Qy 341 AspAsnHisAspAspThrAsnArgIlelleThrValTyrMetGlyLysArgProGluGly 360
D 1962 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 2021
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
```



```
Db 2022 GGAATGCTAGTATCATTTAGCC 2045
RESULT 11
LOCUS I13206 1119 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 30 from patent US 5434073.
ACCESSION I13206
VERSION I13206.1 GI:910554
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 1119)
AUTHORS Dawson,K., Hunter,M.G. and Czaplewski,L.G.
TITLE Fibrinolytic and anti-thrombotic cleavable dimers
JOURNAL Patent: US 5434073-A 30 18-JUL-1995;
FEATURES
  Location/Qualifiers
    source 1..1119
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.35e-138 Length: 1119
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
Db: 6 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x I13206 (1-1119)
Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 7 AGCCAAATTAGTTGTTAGCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 66
Qy 21 LysPheGluLeuAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 67 AAATTTTGAATTTGACCTAACATCAACGACCTGCTCATGGAGGAAAGACAGACGAGC 126
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 127 TTAAGTCCAAATCAAAACCATTTGCTACTGATGATGAGCGCGCATGACATAACTTGA 186
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 187 AAAGCTGACTTACTAAAGCTATTCAAGAACCAATTGATCGTAACTCCACAGTAACGAC 246
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 247 GACTACTTTGAGTCTATTGTTTGCACGGATGCAACCATTTACTGATCGAAACGCAAG 306
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 307 GTCTACTTTGCTGACAAAGATGGTTGCGTAACTTGCACGACCAACCTGTCGAAGAATT 366
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 367 TTGCTAAAGCGGACATGTCGCGCTTAGACCATATAAGAAACCAATACAAATCAAGCG 426
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 427 AAATCTGTTGATGGAATATATCTGTACAGTTTACTTCCCTTTAAACCTGATGACGATTTC 486
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 487 AGACCAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGTGACACCATCACA 546
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 547 TCTCAAGAAATTACTAGCTCAAGCAAAAGCAATTTTAAACAAACCCATCCAGCTATACG 606
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
```

```
Db 607 ATTTATGAACGTCGACTCCTCAATCGTCACATCATGACAATGACATTTTCGTACGATTTTA 666
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 667 CCATGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAGCTTATGATGATCAAT 726
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 727 AAAAAATCTGGTCTGAATGAAGAAATAAACCAACACTGACCTGATCTCTGAGAAATATTAC 786
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 787 GTCCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTC 846
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 847 ACCATCAATACGTTGATGTCAACACCAAGAAATGCTAAAGCCGACGACGCTCTTAACA 906
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 907 GCTAGCGAAAGCTAACTTAGACTTCAGAGATTTATACGATCTCTGTGATAAGGCTAACTA 966
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 967 CTCTACAACAATCTCGATGCTTTTGGTATTATGACTTATACCTTAACCTGAAAAAGTAGAA 1026
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1027 GATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCCGACCCGAAGGA 1086
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1087 GAGAATGCTAGTATCATTTAGCC 1110
RESULT 12
ARI175891 1209 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 11 from patent US 6309873.
ACCESSION ARI175891
VERSION ARI175891.1 GI:17917190
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1209)
AUTHORS Torrens Madrazo,Idel.Carmen., Fuente Garcia,Jdela.,
Ojalvo,A.Garcia., Menendez,A.Seralena., Escalona,E.Pupo.,
Maso,J.Raul.Fernandez. and Gonzalez Griego,Mde.Jesus.
TITLE Streptokinase mutants
JOURNAL Patent: US 6309873-A 11 30-OCT-2001;
FEATURES
  Location/Qualifiers
    source 1..1209
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.47e-138 Length: 1209
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
Db: 6 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x ARI175891 (1-1209)
Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 10 ACCCAATTAGTTGTTAGCGTTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 69
Qy 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
```


Db 70 AAAATTTTGAATTCACCTAACATCAGCACCTGCTCATGAGGAGAAAGACAGAGCAAGGC 129
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 130 TTAAGTCCAAATCAAAACCAATTTGCTACTGATGTAGTGGCGCATGCCATAACTTGAA 189
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 190 AAAGCTGACTTAAAGGCTATTCAAGAACAATTCATCGCTAACCGTCACAGTAAAGC 249
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 250 GACTACTTTGAGGTCAATTCATTTGCAAGCATGCAACCAATTCATGATCGAAACGCAAG 309
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 310 GTCTACTTTGCTGCAAAAGATGGTTGCGTAACCTTGGCGACCAACCTGTCCAAGAAATTC 369
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 370 TTGCTAAGCGGCATGTCGGCTTAGACCATATAAAGAAACCAATACAAATCAAGGC 429
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 430 AAATCTGTTGATGCGAATATACGTACAGTTTACTCCCTTAAACCCCTGATCAGCATTC 489
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 490 AGACCAGGCTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATACA 549
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 550 TCTCAAGAAATTAAGTCTAAGCACAAGCAATTTTAAACAAACCCACCGCTATACG 609
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 610 ATTTATGAACGTGACTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTA 669
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 670 CCAATGGATCAAGAGTTTACTTACCATGTCACAAATTCGGGAAACAAGCTTATGAGATCAAT 729
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 730 AAAAATCTGCTCTGAATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATAC 789
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLysLysLeuPhe 280
Db 790 GTCCTTAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTT 849
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 850 ACCATCAAAATACGTTGATGTCAACCAACGAATTTGCTAAAAAGCGAGCAGCTCTTAA 909
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 910 GCTAGCGAAGCTAACTTAGACTTCAGAGATTTATACGATTCCTCGTGAAGGCTAAACTA 969
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 970 CTCTACAACAATCTCGATGCTTTGTTGTTATATGAGCTATACCTTAACGTGAAAGTAGAG 1029
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1030 GATAATCACGATGACCAACCGTATCATACCGTTTATACCGTTTATGGCAAGCGACCCGAAG 1089
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1090 GAGAAATGCTAGCTATCATTTAGCC 1113
RESULT 13
AX030315
LOCUS AX030315 1209 bp DNA linear PAT 16-SEP-2000

DEFINITION Sequence 11 from Patent EP0985729.
ACCESSION AX030315
VERSION AX030315.1 GI:10190483
KEYWORDS Streptococcus dysgalactiae subsp. equisimilis
SOURCE Streptococcus dysgalactiae subsp. equisimilis
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1
AUTHORS Seralena, M.A., de la Fuente, G.J., Garcia, O.A. and Torrents, M.I.
TITLE Streptokinase mutants
JOURNAL Patent: EP 0985729-A 11 15-MAR-2000;
CIGB (CU)
FEATURES
Location/Qualifiers
source 1..1209
/organism="Streptococcus dysgalactiae subsp. equisimilis"
/mol_type="unassigned DNA"
/sub_species="equisimilis"
/db_xref="taxon:119602"
misc_feature 1..1209
/note="Nucleotide sequence encoding a streptokinase SKC-2
fragment (40-1245)"
ORIGIN
Alignment Scores:
Pred. No.: 1,47e-138 Length: 1209
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
DB: Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x AX030315 (1-1209)
Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 10 AGCCAAATTAAGTTAGTGTGCTGCTACTGTTGAGGGGACGAATCAAGACATTAAGTCTT 69
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 70 AAAATTTTGAATTCACCTAACATCAGCACCTGCTCATGAGGAGAAAGACAGAGCAAGC 129
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 130 TTAAGTCCAAATCAAAACCAATTTGCTACTGATGTAGTGGCGCATGCCATAACTTGAA 189
Qy 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 190 AAAGCTGACTTAAAGGCTATTCAAGAACAATTCATCGCTAACCGTCACAGTAAAGC 249
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 250 GACTACTTTGAGGTCAATTCATTTGCAAGCATGCAACCAATTCATGATCGAAACGCAAG 309
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 310 GTCTACTTTGCTGCAAAAGATGGTTGCGTAACCTTGGCGACCAACCTGTCCAAGAAATTC 369
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 370 TTGCTAAGCGGCATGTCGGCTTAGACCATATAAAGAAACCAATACAAATCAAGGC 429
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 430 AAATCTGTTGATGCGAATATACGTACAGTTTACTCCCTTAAACCCCTGATCAGCATTC 489
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 490 AGACCAGGCTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATACA 549
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 550 TCTCAAGAAATTAAGTCTAAGCACAAGCAATTTTAAACAAACCCACCGCTATACG 609


```
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 826 GTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTC 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAAAATACGTTGATGTCACACCAACGAATTCGTAAGGCGAGCAGCTCTTAAACA 945
Qy 301 AlaSerGluAArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTACGCAACGAACTTAGACTTCAGAGATTTATACGATCCCTCGTGAAGGCTAAACTA 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTCTACACAATCTCGATGCTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GATAATCAGGATGACACCAACCGTATCATACCGTTTATATGGCAAGCGCCCGAAGGA 1125
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGAAATGCTAGCTATCATTTAGCC 1149

RESULT 15
AR175892
LOCUS AR175892 1245 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 12 from patent US 6309873.
ACCESSION AR175892
VERSION AR175892.1 GI:17917191
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE
1 (bases 1 to 1245)
Torrens Madrazo,Idel. Carmen., Fuente Garcia,Jdela.,
Ojalvo,A.Garcia., Menendez,A.Seralena., Escalona,E.Pupo.,
Maso,J.Raul.Fernandez. and Gonzalez Griego,Mde.Jesus.
TITLE Streptokinase mutants
JOURNAL Patent: US 6309873-A 12 30-OCT-2001;
FEATURES
Location/Qualifiers
1..1245
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 1.52e-138 Length: 1245
Score: 1875.00 Matches: 363
Percent Similarity: 98.91% Conservative: 1
Best Local Similarity: 98.64% Mismatches: 4
Query Match: 98.63% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AR175892 (1-1245)

Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 10 AGCCAATTTAGTTGTAGCGTTTCTGCTACTGTTGAGGGACGAATCAAGACATTAGCTTT 69
Qy 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 70 AAAATTTTGGAAATTTGACCTAATACATCACCACTGCTCATGGAGGAAAGACAGAGCAGGC 129
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValaMetSerHisLysLeuGlu 60
Db 130 TTAAGTCCAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGCCACATAAATTTGAA 189
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 190 AAAGCTGACTTACTAAAGCTATTCAAGAACAAATTCATCGCTAAACGCTCCACAGTAACGAC 249
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
```

```
Db 250 GACTACTTTTCAGTCAATTTGCAAGCGATGCAACCACTTACTGATCGAAACGGCAAG 309
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 310 GTCTACTTTTCTGCTGACAAAGATGGTTTCGGTAACCTTCCGACCCCAACCTGTCGAAGATTT 369
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 370 TTGCTAAGCGGACATGTGGCGTTAGACCATATAAGAAAAAACCAATACAAATCAAGCG 429
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 430 AAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTTC 489
Qy 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 490 AGACCAAGGCTCAAGATATACTAAGCTATTGAAACACTAGCTATCGTGACACCATCACA 549
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 550 TCTCAAGAAATTACTAGCTCAAGCACAAGCATTTTAAACAAACCCACCAGGCTATACG 609
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 610 ATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTA 669
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 670 CCAATGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAAGCTTATGAGATCAAT 729
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 730 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 789
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 790 GTCCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTC 849
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 850 ACCATCAAAATACGTTGATGTCACACCAACGAATTCGTAAGGCGAGCAGCTCTTAAACA 909
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 910 GCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCCTCGTGAATAGGCTAAACTA 969
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 970 CTCTACAAACAACTCGATGCTTTTGGTATTATGACTATACCTTAACCTGGAAGTAGAG 1029
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1030 GATAATCAGGATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGCCCGAAGGA 1089
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1090 GAGAAATGCTAGCTATCATTTAGCC 1113
```

Search completed: November 6, 2004, 02:55:01
Job time : 5212.77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 5, 2004, 21:47:11 ; Search time 589.178 Seconds
(without alignments)
3278.783 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901

Sequence: 1 SQLVVSAGTVEGTNQDISL.....IITVYMKRPEGENASYHLA 368

Scoring table:-

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09940235/runat_03112004_174038_11244/app_query.fasta_1.1045
-DB=N Geneseq_23Sep04 -QFT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235@CGN_1_1063@runat_03112004_174038_11244 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No	Score	Query Match	Length	ID	Description
1	1901	100.0	1242	2	Aax80492 Streptoco
2	1901	100.0	1242	2	Aax16632 Streptoco
3	1901	100.0	1245	3	Aaa37633 S. equisi
4	1901	100.0	1254	6	Aba05546 Streptoki
5	1901	100.0	1262	2	Aaq10230 Synthetic
6	1901	100.0	1327	3	Aaa37628 Streptoki

7	1901	100.0	1377	3	AAA37622	Aaa37622 Streptoki
8	1901	100.0	1541	3	AAA37644	Aaa37644 Chimeric
9	1901	100.0	1661	3	AAA37637	Aaa37637 Chimeric
10	1901	100.0	1782	3	AAA37642	Aaa37642 Chimeric
11	1901	100.0	2096	3	AAA37643	Aaa37643 Chimeric
12	1901	100.0	2385	2	AAX80497	Aax80497 Streptoki
13	1901	100.0	8893	6	ABA05547	Aba05547 Maxadlian
14	1898	99.8	2030	2	AAQ11651	Aaq11651 FB-FB-SK
15	1897	99.8	1242	2	AAQ11663	Aax16633 Streptoco
16	1875	98.6	1119	2	AAQ12159	Aaq12159 Truncated
17	1875	98.6	1209	3	Aaz99249	Aaz99249 DNA encod
18	1875	98.6	1245	3	Aaz99250	Aaz99250 DNA encod
19	1875	98.6	1335	2	AAQ12156	Aaq12156 Streptoki
20	1875	98.6	1458	2	AAQ12162	Aaq12162 Factor xa
21	1875	98.6	1467	2	AAQ12490	Aaq12490 Factor xa
22	1875	98.6	1512	2	AAQ12158	Aaq12158 Streptoki
23	1875	98.6	2253	2	AAQ12161	Aaq12161 Met-core
24	1875	98.6	2589	2	AAQ12160	Aaq12160 OmpAL str
25	1875	98.6	7057	12	ADM01294	Adm01294 Plasmid p
26	1871.5	98.4	2566	2	AAT77778	Aat77778 Coding se
27	1871	98.4	1245	2	AAQ05603	Aaq05603 SKC-2 str
28	1859	97.8	1473	2	AAQ05603	Aaq05603 Streptoki
29	1857	97.7	1407	1	AAN70106	Aan70106 DNA encod
30	1836	96.6	1323	2	AAT29961	Aat29961 Vector ps
31	1819	95.7	1158	3	Aaz99252	Aaz99252 DNA encod
32	1817	95.6	1242	5	AAF82144	Aaf82144 Mutant st
33	1815	95.5	1122	3	Aaz99251	Aaz99251 DNA encod
34	1786	94.0	2568	1	AAN50493	Aan50493 Sequence
35	1741	91.6	1320	6	ABN70192	Abn70192 Streptoco
36	1726	90.8	1473	2	AAQ05604	Aaq05604 Streptoki
37	1686.5	88.7	2208	2	AAX83589	Aax83589 Recombina
38	1686	88.7	1068	2	AAX80493	Aax80493 Recombina
39	1584	83.3	1245	10	ADP48644	Adf48644 Streptoco
40	785	41.3	450	2	AAX80494	Aax80494 Recombina
41	275	14.5	1181	2	AAV84164	Aav84164 Streptoco
42	267	14.0	1180	2	AAV84163	Aav84163 Streptoco
43	138.5	7.3	7320	2	AAT68081	Aat68081 H. pylori
44	127	6.7	110000	6	ABN71527	Continuation (11 o
45	127	6.7	110000	6	ABN71527_11	Continuation (12 o

ALIGNMENTS

RESULT 1

AAX80492
ID AAX80492 standard; cDNA; 1242 BP.

XX
AC AAX80492;

XX
DT 17-OCT-2003 (revised)

DT 26-AUG-1999 (first entry)

XX
DE Streptococcus equisimilis native streptokinase encoding cDNA.

XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;

KW NSK; RSK; bacterial; blood clot; thrombotic condition;

KW myocardial infarction; venous thrombosis; pulmonary embolism;

KW cerebral thrombosis; graft thrombosis; arterial thrombosis; ss.

XX
OS Streptococcus dysgalactiae subsp. equisimilis.

XX
PN WO9931247-A1.

XX
PD 24-JUN-1999.

XX
PF 15-DEC-1998; 98WO-US026694.

XX
PR 15-DEC-1997; 97US-0069497P.

XX
PA (HARD) HARVARD COLLEGE.

XX
PI Reed GL;

XX
PI Reed GL;

DR WPI; 1999-395183/33.
DR P-PSDB; AAY24794.
XX
XX N-terminally deleted streptokinase.
XX
XX Claim 44; Page 58-60; 73pp; English.
XX
XX The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein; a
CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising a
CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
CC blood clots in patients with a thrombotic condition, e.g. myocardial
CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
CC graft thrombosis and arterial thrombosis. The modified streptokinase can
CC also be used in non-human mammals. Streptokinase activation of
CC plasminogen is at least 10-fold, preferably 100-fold greater in the
CC presence of fibrin than in the absence of fibrin. The modified
CC streptokinase has at least one amino acid substitution that inactivates a
CC substrate site for proteolytic cleavage. This reduces the rate of
CC degradation of the streptokinase at least two-fold. The present sequence
CC encodes native streptokinase (nsK). (Updated on 17-Oct-2003 to
CC standardise OS field)
XX
SQ Sequence 1242 BP; 424 A; 267 C; 237 G; 314 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,12e-176 Length: 1242
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAX80492 (1-1242)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerIleu 20
Db 46 AGCCAAATTAGTTGTTAGCGTTCTGCTACTGTTGAGGGGACCAATCAAGACATTATGCTT 105
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 106 AAATTTTGAATCGATTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGGC 165
Qy 41 LeuSerProLysSerIysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 TTAAGTCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCATGTCAATAAATTTGAG 225
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 226 AAAGTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAAACGAC 285
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGGTCAATGATTTTGCACGATGCAACCATTTACTGATCGAAACGGCAAG 345
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 346 GTCTACTTTGCTGACAAAGATGTTTCGGTAACTTGGCCGACCAACCTGTCCTCAAGAAATTT 405
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 TTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG 465
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 466 AAATCTGTTGATGTGAATATATCTGTACAGTTTACTCCCTTTAAACCTGATGACGATTTTC 525

Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 AGACCAAGGTTCTCAAGATACTAAGCTATTGAAAAACACTAGCTATCGGTACACCATCACA 585
Qy 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAAGAAATTACTAGCTCAAGCACAAGCAATTTTAAACAAAAACCAACCCAGCTATACG 645
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 646 ATTTATGAACGAGTCTCTCAATCGCTCACTCATGACCAATGACATTTTCCGTACGATTTA 705
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 706 CCAATGATCAAGAGTTTACTTACCGTGTAAANAATCGGGAACAGCTTATAGGATCAAT 765
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyrTyr 260
Db 766 AAAAATCTGGTCTGAATGAAGAATAAAACACACTGACCTGATCTCTGAGAAATATTAC 825
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLysLeuPhe 280
Db 826 GTCTCTTAAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTT 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAAAATAGTGTGATGTCATACCAACGAATTTGCTAAAAGTAGGAGCGCTCTTAACA 945
Qy 301 AlaserGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTAGGGAACGTAACCTTAGACTTCAGAGATTTATACGATCTCTGATGAAGGCTAACTA 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTCTACCAACATCTCGATGCTTTTGGTATTATGAGCTATACCTTAACTGGAAGTAGAG 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGlyGly 360
Db 1066 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCACCAGGA 1125
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGATGCTAGCTATCATTTAGCC 1149
RESULT 2
ID AAX16632
XX AAX16632 standard; DNA; 1242 BP.
AC AAX16632;
XX
DT 17-OCT-2003 (revised)
DT 04-MAY-1999 (first entry)
XX
DE Streptococcus equisimilis native streptokinase encoding DNA.
XX
KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance; ds.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
XX Key Location/Qualifiers
FT CDS 1..1242
FT /tag= a
FT /transl_except= (pos:40..42,aa:Asn)
FT FT
XX US5876999-A.
XX
PD 02-MAR-1999.

PR 24-DEC-1998; 98IN-DE003825.
XX (COUL) CSIR COUNCIL SCI IND RES.
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI P-PSDB; AAY90282.
XX WPI; 2000-516032/47.
DR P-PSDB; AAY90282.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibrinectin.
XX
XX Example 3; Fig 3; 59pp; English.
XX This sequence represents the human Streptococcus equisimilis
CC streptokinase coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibrinectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 1245 BP; 426 A; 267 C; 237 G; 315 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2, 13e-176 Length: 1245
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAA37633 (1-1245)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 46 AGCCAAATTTAGTTGTTAGCGTTCCTGCTACTGTTGAGGGGACGAATCAAGACATTAGTCTT 105
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 106 AAAATTTTGAATCGATTAACATCACGACCTGCTCATGAGGAAAGACAGACGAAGC 165
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 TTAAGTCCAAATCAAAACCAATTTCTGCTAGTAGTGGCGCGATGTCACATAACTTAG 225
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 226 AAAGCTGACTTAAGGCTATTCAAGACAAATTCATGCTTAACGTCACAGTAACGAC 285
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGGTCAATGATTTTCAAGCGATGCAACCATTAATTAATCAAGCAAG 345
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnPhe 120

Db 346 GTCTACTTTTCTGACAAAGATGGTTTCGTAAACCTTCGCCGACCAACCTCTCCAAGAAATTT 405
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 TTGCTAAGCGGACATGTGCGGTAGACCATATAAAGAAACCAATATACAAACCAAGCG 465
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 466 AAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCTTGATGAGATTTC 525
Qy 161 ArgProGlyLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 AGACACAGGCTCAAGACATATAAGCTATTCAAAACACTAGCTATATCGGTGACCATCACA 585
Qy 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAACCAACCCAGGCTATACG 645
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 646 ATTTATGAACGCTGACTCTCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTA 705
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 706 CCAATGGATCAAGAGTTTACTTTACCGTGTAAAAATCGGAAACAAGCTTATAGGATCAAT 765
Qy 241 LysLysSerGlyLeuAsnGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 766 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 825
Qy 261 ValLeuLysLysGlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 826 GTCTTAAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAACACTGTTTC 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAAAATACGTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAAACA 945
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTAGCGAAACGTAACCTTAGACTTCAGAGATTTTATACGATCCCTCGTGAAGGCTAAACTA 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GATATCACGATGACACCAACCGTATACCGTTTATATGCGCAAGCGACCCGAAGGA 1125
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGAATGCTAGCTATCATTTAGCC 1149
RESULT 4
ABR05546
ID ABR05546 standard; cDNA; 1254 BP.
XX
AC ABA05546;
DT 26-FEB-2002 (first entry)
XX
DE Streptokinase cDNA.
XX Streptokinase; cerebroprotective; cardiant; gene therapy; fusion;
KW vasodilator; thrombolytic; angina; myocardial infarction; stroke;
XX gene therapy; maxadilan; ss.
OS Unidentified.
XX
PN W0200185100-A2.
XX
PD 15-NOV-2001.

10-MAY-2001; 2001WO-US015209.
11-MAY-2000; 2000US-00569920.
(GEHO) GEN HOSPITAL CORP.
Reddy VB, Lerner E;
WPI; 2002-062184/08.
New fusion protein or conjugate, useful for treating unstable angina, acute myocardial infarction or stroke, comprises a vasodilator polypeptide and a thrombolytic polypeptide, or active fragments of the polypeptides.
Example 1; Fig 2; 37pp; English.
The invention relates to a fusion protein or a conjugate comprising a vasodilator polypeptide, or its active fragment, and a thrombolytic polypeptide or its active fragment. The protein is useful for treating a subject suffering from a partially or totally occluded blood vessel, causing unstable angina, acute myocardial infarction or stroke. The polynucleotide encoding the polypeptide is useful in gene therapy. The vasodilatory action of the protein allows for the use of lower doses of the thrombolytic while maintaining the clot dissolving effectiveness of the thrombolytic, and the use of lower doses of the thrombolytic reduces associated side effects. The present sequence is the streptokinase cDNA used in the construction of a Maxadilan-Streptokinase fusion protein. Maxadilan is a vasodilator peptide produced by the salivary gland of the New World sand fly
Sequence 1254 BP; 426 A; 271 C; 241 G; 316 T; 0 U; 0 Other;

Db	472	AAATCTGTGTGATGGGAATATAC	TGTA	CGTTTACTCCCTTAAACCCCTGATGACGAT	TTC	531
Qy	161	ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr	180			
Db	532	AGACAGGCTCTCAAGAATAC	CTAAGCTAT	TGNAACACATAGCTATCGGTGACACCAT	CACA	591
Qy	181	SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr	200			
Db	592	TCTCAAGAAATTTACTAGCTCAAGCACAAAGCAATTTTAAACAAAAACCCAGGCTATACG	651			
Qy	201	IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu	220			
Db	652	ATTTATGAACGTGACTCTCTCAATCGTCATCAATGACATGACATTTTCCGTACGATTTTA	711			
Qy	221	ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn	240			
Db	712	CCAATGATCAAGAGTTTACTTACCGTGTAAANAATCGGNAACAAGCTTATAGGATCAAT	771			
Qy	241	LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr	260			
Db	772	AAAAAATCTGCTCTGAATGAAGAAATAAACAACTGACCTGATCTCTGAGAAATATTAC	831			
Qy	261	ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe	280			
Db	832	GTCCTTAAAAAGGGGAAAAACCGCATATGATCCCTTTGATCGACGTCACTTGAAACTGTTTC	891			
Qy	281	ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr	300			
Db	892	ACCATCAAAATACGTTGATGTCGATACCACGAATGCTTAAAGTAGAGAGCTCTTAACA	951			
Qy	301	AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu	320			
Db	952	GCTAGCGAACGTAACCTTAGACTTCAGAGATTTTATACGATCCTCGTGATAAGGCTAAACTA	1011			
Qy	321	LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu	340			
Db	1012	CTCTACAAACATCTCGATGCTTTTGGTATTATGGACTATACCTTTAACTGGAAAGTAGAG	1071			
Qy	341	AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly	360			
Db	1072	GATTAATCAGATGACACCAACCGTATCATTAACCGTTTATATGGCAAGGACCCGAAGGA	1131			
Qy	361	GluAsnAlaSerTyrHisLeuAla	368			
Db	1132	GAGAAATGCTAGCTATCATTTAGCC	1155			
RESULT 5						
AAQ10230						
ID	AAQ10230	standard; DNA; 1262 BP.				
XX	AAQ10230;					
AC	AAQ10230;					
XX	28-MAR-1991	(first entry)				
DT						
XX	Synthetic	Streptokinase gene.				
DE						
XX		streptokinase; thrombolytic agent; myocardial infarction; ds.				
KW						
XX						
OS		Synthetic.				
XX						
Key		Location/Qualifiers				
FT	15..1256					
CDS	/*tag= a					
FT	/product= "streptokinase"					
XX						
XX	EP407942-A.					
XX						
PD	16-JAN-1991.					
XX						
PF	11-JUL-1989;	89JP-00179432.				
XX						
XX	11-JUL-1989;	89JP-00179432.				
XX						

PR 27-NOV-1989; 89JP-00307957.
PR 11-APR-1990; 90JP-00096830.
XX
PA (SAXA) OTSUKA PHARM FACTOR.
XX
XX Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;
XX P-PSDB; AAR10194.
DR WPI; 1991-016179/03.
DR P-PSDB; AAR10194.
XX
PT Synthetic gene encoding streptokinase - scale, high purity prodn. of
PT streptokinase used as a thrombolytic agent.
XX
XX Claim 3; Page 56; 76pp; English.
XX
CC The 5' end of the coding strand overhangs the 3' end of the complementary
CC strand by 4 bases; the 5' end of the complementary strand overhangs the
CC 3' end of the sense strand by AGCT. The sequence encoding streptokinase
CC was first divided into 52 oligonucleotide fragments of 43-56 bases each.
CC The individual fragments were synthesised by solid-phase beta-cyanomethyl
CC phosphoramidite method, phosphate groups added and the fragments ligated
CC to obtain pSKK. This was inserted into expression vector pKTN2 and used
CC to transform E.coli JM109 which was cultured to produce streptokinase.
CC See also AAR10195-R10200
XX
SQ Sequence 1262 BP; 365 A; 350 C; 262 G; 285 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.16e-176 Length: 1262
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAQ10230 (1-1262)

QY 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
DB 60 TCCCAAGCTGGTGTTCGGTACCTGGCACTGTGAAGGTACTAAACGAGGACATCTCTCTG 119
QY 21 LysPheGluLeuAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
DB 120 AAAATTTTCGAAATCGACTGACCTCTGCTCGCGCCATGGTGTAAACCGNACAGGGC 179
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
DB 180 CTGTCCTCCGAAATCTAAACCGTTCGCTACTGACTCTGGCGCTATGCTCATAACTCGAG 239
QY 61 LysAlaAspLeuLysAlaAlaGlnGluGlnLeuLeuAlaAsnValHisSerAsnAsp 80
DB 240 AAGGAGATCTGCTGAAAGCAATCCAGGAACAGCTGATCGCTAAACGTATCTAAACGAC 299
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
DB 300 GACTACTTTAGGTAATCGACTTCGCTAGCGAGCTTACTATCACCAGCGTAAACGCAAA 359
QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
DB 360 GTATACTTCGCTGACAAAGACGGTTCGTAACTCTTCCGACTCAACCGGTACAGGAATTT 419
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
DB 420 CTGCTGTCTGGCCATGTACCGCTTCGCCCGGTAAAGAAAAACCGATCCAGAACCAAGGCT 479
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
DB 480 AAATCTGTGTAGGTAGATACACCGTTCACCGCTGAAACCGATCCAGACGATGACTTC 539
QY 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
DB 540 CGCCCGGCTCGAAGACACTAAACTGCTGAAACCTCGCTATCGGTGACACCATCACT 599

QY 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
DB 600 TCTCAGAGCTCTGCTGCTCAGGCACAGTCTATCTGAACAAAAACCATCCGGCTACACT 659
QY 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
DB 660 ATCTAGCAACGCGACTCTTCCATCGTAACCCATGACACCATCTTCCGTACCATCTCTG 719
QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
DB 720 CCATGAGCAGAGAAATTTACTTACCGTGTAAACACCGCAACGCTTACCGTATCAAT 779
QY 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
DB 780 AAAAATCCGCTCTGAATGAGAGATTACACACTGACTGATCTCTGAAGTACTTAC 839
QY 261 ValLeuLysLysGlyGlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
DB 840 GTACTGAAAAAAGGTGAGAAGCGGTATGACCCGTTCCGATCGTTCTCATCTGAAACTGTT 899
QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
DB 900 ACCATCAATACGTTGACGTCGATACCAACGAATTACTGAAGTCTGAGCAGCTGCTGACC 959
QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
DB 960 GCTTCGAAACGTAATCTGGACTTCCGATCTGTGACGCCGCTGACAAAGTAAACTG 1019
QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
DB 1020 CTGTACAACAACCTGGATGCTTTCGGTATCATGGACTACACCTGACTGTTAAAGTAGAA 1079
QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
DB 1080 GACACCATGACGACACCAACCGTATCATCACCGTATATACCGTATATGGGCAACGTC 1139
QY 361 GluAsnAlaSerTyrHisLeuAla 368
DB 1140 GAAATGCACTTCTTACCATCTGGCA 1163
RESULT 6
AAA37628
ID AAA37628 standard; DNA; 1327 BP.
XX AAA37628;
AC
XX
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE Streptokinase-NTR gene.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibrinectin; thrombolytic therapy;
XX cardiovascular disorder; ss.
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN EP1024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
XX (COUL) CSIR COUNCIL SCI IND RES.
XX
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
XX WPI; 2000-516032/47.
XX
PT Hybrid streptokinase-fibrin binding domain polypeptides useful for

thrombolytic therapy comprises a streptokinase fused with fibrin binding domains of human fibronectin.

Example 1; Fig 14; 58pp; English.

This sequence represents a streptokinase-NTR (SK-NTR) gene (where NTR stands for N-terminally repaired with native sequence). The invention relates to a hybrid plasminogen activator (PA) comprises a polypeptide fusion between streptokinase (SK), which are capable of plasminogen (PG) activation, and fibrin binding regions of human fibronectin, which are from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the ability to bind with fibrin independently and also characteristically retains a PG activation ability which becomes evident only after a pronounced duration, or lag, after exposure of the PA to a suitable animal or human PG. The hybrid streptokinase-fibrin binding domain polypeptides are useful in thrombolytic therapy for various kinds of cardiovascular disorders. The hybrids have enhanced fibrin selectivity as well as kinetics of plasminogen activation that are distinct from that of natural streptokinase in being characterised by a temporary delay, or lag of several minutes in the natural rate of the catalytic conversion of plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins can bind tightly with fibrin in blood clots soon after introduction into the vascular system without significantly activating the circulating blood plasminogen to plasmin, thus aiding in the localisation of the plasminogen activation process to the site of pathological thrombus. This overcomes systemic plasminogen activation encountered during clinical use of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 1327 BP; 456 A; 279 C; 250 G; 342 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,3e-176 Length: 1327
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAA37628 (1-1327)

Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 128 AGCCAAATGGTTGTAGCGTGTCTGTTGTTGAGGAGCAAGTCAAGACATTAGTCTT 187
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 188 AAATTTTGGAAATCGATTAAACATCAGCAGCTGCTCATGGAGGAAAGACAGAGCAAGC 247
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValaMetSerHisLysLeuGlu 60
Db 248 TTAAGTCCAAATCAAAACATTTGCTACTGATGATGGCGGATGTCACATAAATTTGAG 307
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 308 AAAGCTGACTTACTAAAGGCTATTCAAGAACCAATTGCTGCTAAAGTCCACAGTAAAGC 367
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 368 GACTACTTTGAGGTCAATTTGCAAGCGATGCAACCATTTACTGATGAAACGCGAAG 427
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 428 GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCCAACCTGTCCAAGAAATTT 487
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 488 TTGCTAAGCGGACATGTCGCGTTAGACCATATAAGAGAAACCAATACAAAACCAAGCG 547
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 548 AAATCTGTGTGATGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTTC 607
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180

Db 608 AGACCAAGGCTCAAGATACCTAAGCTATTGAAACACTAGCTATCGGTGACCATCACA 667
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 668 TCTCAAGAATTACTAGCTCAAGCACAAGCATTTTAAACAAAACCAACCCAGGCTATACG 727
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 728 ATTTATGAACGTCGACTCCTCAATCGTCACATCATGACAATGACATTTTCCGTCAGATTTA 787
Qy 221 ProMetAspGlnGluPheThrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 788 CCAATGGATCAAGAGTTTACTTCCGCTGTTTAAATAATCGGAACAAGCTTATAGGATCAAT 847
Qy 241 LysLysSerGlyLeuAsnGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 848 AAAAATCTGCTGTAATGAAGAATAAACAACACTGACCTGATCTCTGAGAATATTAC 907
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 908 GTCCTTAAAAAGGGGAAAGCGGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTC 967
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 968 ACCATCAAAATACGTTGATGTGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACA 1027
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 1028 GCTAGCGAAGCTAATCTAGACTTCAGAGATTTATACGATCCTCGTATAAGGCTAAACTA 1087
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1088 CTCTACAACATCTCGATGCTTTTGGTATTATGACTATACCTTAACCTGAAAAGTAGAG 1147
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1148 GATAATCAGCATGACACCAACCGTATCATAAACCGTTTATATGGCAAGCGACCCGAAGGA 1207
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1208 GAGAATGCTAGCTATCATTTAGCC 1231
RESULT 7
AAA37622
ID AAA37622 standard; DNA; 1377 BP.
XX
AC AAA37622;
XX
XX 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE Streptokinase-NTR gene.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
XX
XX cardiovascular disorder; ss.
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
PN EP1024192-A2.
XX
XX 02-AUG-2000.
XX
XX 23-DEC-1999; 99EP-00310541.
XX
XX 24-DEC-1998; 98IN-DE003825.
XX
XX (COUL) CSIR COUNCIL SCI IND RES.
XX
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX

(COUL) CSIR COUNCIL SCI IND RES.

Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V; Yadav M;

WPI; 2000-516032/47.

Hybrid streptokinase-fibrin binding domain polypeptides useful for thrombolytic therapy comprises a streptokinase fused with fibrin binding domains of human fibronectin.

Disclosure; Fig 17b; 58pp; English.

This sequence represents a chimeric streptokinase-fibrin binding domain (SK-FBD) protein coding sequence. The invention relates to a hybrid plasminogen activator (PA) comprises a polypeptide fusion between streptokinase (SK), which are capable of plasminogen (PG) activation, and fibrin binding regions of human fibrinectin, which are from fibrin binding domains (FBD) 4 and 5 of 1 and 2. The hybrid PA possesses the ability to bind with fibrin independently and also characteristically retains a PG activation ability which becomes evident only after a pronounced duration, or lag, after exposure of the PA to a suitable animal or human PG. The hybrid streptokinase-fibrin binding domain polypeptides are useful in thrombolytic therapy for various kinds of cardiovascular disorders. The hybrids have enhanced fibrin selectivity as well as kinetics of plasminogen activation that are distinct from that of natural streptokinase in being characterised by a temporary delay, or lag of several minutes in the natural rate of the catalytic conversion of plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins can bind tightly with fibrin in blood clots soon after introduction into the vascular system without significantly activating the circulating blood plasminogen to plasmin, thus aiding in the localisation of the plasminogen activation process to the site of pathological thrombus. This overcomes systemic plasminogen activation encountered during clinical use of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

XX 23-DEC-1999; 99EP-00310541.
PF CC
PR 24-DEC-1998; 98IN-DE003825.
XX (COUL) CSIR COUNCIL SCI IND RES.
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
PI WPI; 2000-516032/47.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
PT domains of human fibronectin.
XX Example 5; Fig 19b; 58pp; English.
PS This sequence represents a chimeric streptokinase-fibrin binding domain
XX (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
XX SK Sequence 1661 BP; 529 A; 370 C; 357 G; 405 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,05e-176 Length: 1661
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAA37637 (1-1661)

Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeu 20
Db 228 AGCCAAATGGTGTAGCGTGTCTGCTGTTGAGGGGACCAATCAAGACATAGTCTT 287

Qy 21 LysPheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 288 AAAATTTTGAATCGATTAACATCACGACCTGCTCATGGAGGAAAGACAGCAAGGC 347

Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 348 TTAAGTCCAAATCAAAACCAATTTGCTGATGATGAGCGCGCATGTCACATAAATAGAG 407

Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 408 AAAGCTGACTTAAAGGCTATTCAAGNACAAATGATCGCTAAAGTCCACAGTAACGAC 467

Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgGlnGlyLys 100
Db 468 GACTACTTTGAGTTCATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCGAAG 527

Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 528 GTCTACTTTGCTGCACAAAGATGTTCCGTTAAACCTTCCGACCCCAACCTCTCCAAAGATTT 587

Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnGlnAla 140
Db 588 TTGCTTAAGCGGACATGTGCGGTTAGACCATATAAAGNAAAAACCAATACAAAACCAAGCG 647

Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 648 AAATCTGTTGATGTGGAATATACTGTACAGTTACTCCCTTAAACCCCTGATGACGATTC 707

Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 708 AGACCAGGTTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 767

Qy 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 768 TCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCAGGCTATACG 827

Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 828 ATTATGAACGAGTACTCTCAATCGTCACTCATGACATGACATTTTCGTACGATTTTA 887

Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 888 CCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAAT 947

Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 948 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1007

Qy 261 ValLeuLysGlyGlyLysProTyrAspProPheAspAspSerHisLeuLysLeuPhe 280
Db 1008 GTCTTTAAAAAAGGGGAAAGCGGTAGATCCCTTTGATCGCAGTCCTTGAAACTGTTC 1067

Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 1068 ACCATCAATAGCTTGATGTCGATACCAAGCAATTGCTAAAAGTCGACGAGCTCTTAAACA 1127

Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 1128 GCTAGCGAAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGATAGGCTAAACTA 1187

Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1188 CTCACAAACATCTCGATGCTTTTGGTATTTATGGACTATACCTTAACTGGAAAGTAGAG 1247

Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1248 GATNATCAGATGACACCAACCGGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 1307

Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1308 GAGAAATGCTAGCTATCATTTAGCC 1331

RESULT 10
AAA37642
ID AAA37642 standard; DNA; 1782 BP.
XX
AC AAA37642;
XX
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE Chimeric SK-FBD coding sequence.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.

OS Chimeric.
 XX EP1024192-A2.
 XX PD 02-AUG-2000.
 XX PF 23-DEC-1999; 99EP-00310541.
 XX XX 24-DEC-1998; 98IN-DE003825.
 XX (COUL) CSIR COUNCIL SCI IND RES.
 XX PA Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX WPI; 2000-516032/47.
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX Example 5; Fig 21b; 58pp; English.
 XX This sequence represents a chimeric streptokinase-fibrin binding domain
 CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)
 XX SQ Sequence 1782 BP; 573 A; 407 C; 377 G; 425 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 3.33e-176 Length: 1782
 Score: 1901.00 Matches: 368
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-358) x AAA37642 (1-1782)

QY 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
 Db AGCCAAATGGTTGTTAGCGTGTGCTACTGTTGGGGGACGAATCAAGACATTAGTCTT 642
 QY 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
 Db 643 AAAATTTTGAATCGATTAACATCAGCACCTGCTCTATGATGGCGGCAAGACAGAGC 702
 QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValaMetSerHisLysLeuGlu 60
 Db 703 TTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGCATGTCACATAAATTTGAG 762
 QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
 Db 763 AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGTAACTCCACAGTACAGC 822

QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
 Db 823 GACTACTTTGAGGTCAATTGATTTGCAAGCGATGCAACCACTTACTGATGAAACGGCAAG 882
 QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
 Db 883 GTCTACTTTGCTGACAAAGATGGTTCCGTAACTTCCGCGACCCCAACCTGCTCCAGAAATTT 942
 QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
 Db 943 TTGCTAAGCGGACATGTGCGCGTTAGACCATATATAAGAAAAACCAATACAAACCAAGCG 1002
 QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
 Db 1003 AAATCTGTTGATGTGGAATATACTGTACAGTTACTCCCTTAAACCCCTGATGACGATTC 1062
 QY 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
 Db 1063 AGACCAAGGTCTCAAGATACTAAAGTATTGAAACACTAGCTATCGGTGACACCATCACA 1122
 QY 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
 Db 1123 TCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAAACCAACCCAGGCTATACG 1182
 QY 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
 Db 1183 ATTTATGAACGTGACTCTCTCAATCGTCACATCATGACCAATGACATTTTCCGACGATTTA 1242
 QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaIleArgIleAsn 240
 Db 1243 CCAATGATCAAGAGTTTACTTACCGGTGTTAAAAATCGGAAACCAAGCTTATAGGATCAAT 1302
 QY 241 LysLysSerGlyLeuAsnGluIleAsnAsnThrAspLeuIleSerGlyLysTyrThr 260
 Db 1303 AAAAAATCTGGTCTGAATGAAGAAATAAAACAACACTGACCTGATCTCTGAGAAATATATAC 1362
 QY 261 ValLeuLysLysGlyGlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
 Db 1363 GTCTTAAAAAAGGGGNAAGCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTTC 1422
 QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
 Db 1423 ACCATCAATACGTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACA 1482
 QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
 Db 1483 GCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTATAGGGCTAAACTA 1542
 QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
 Db 1543 CTCTACAAACAACTCGATGCTTTTGGTATTATGACTATACCTTAACCTGAAAGATAGAG 1602
 QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
 Db 1603 GATAATCACGATGACACCAACCGCTATCATAAACCGTTTATATGGCAAGCGACCCGAGGA 1662
 QY 361 GluAsnAlaSerTyrHisLeuAla 368
 Db 1663 GAGAATGCTAGCTATCATTTAGCC 1686

RESULT 11

AAA37643
 ID AAA37643 standard; DNA; 2096 BP.

XX AC AAA37643;

XX DT 15-SEP-2003 (revised)

XX DT 13-OCT-2000 (first entry)

XX Chimeric SK-FBD coding sequence.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

plasminogen; human; fibrinectin; thrombolytic therapy; cardiovascular disorder; ss.

Streptococcus dysgalactiae subsp. equisimilis.
Homo sapiens.
Chimeric.

EP1024192-A2.

02-AUG-2000.

23-DEC-1999; 99EP-00310541.

24-DEC-1998; 98IN-DE003825.

(COUL) CSIR COUNCIL SCI IND RES.

Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V; Yadav M;

WPI; 2000-516032/47.

Hybrid streptokinase-fibrin binding domain polypeptides useful for thrombolytic therapy comprises a streptokinase fused with fibrin binding domains of human fibrinectin.

Example 6; Fig 22b; 58pp; English.

This sequence represents a chimeric streptokinase-fibrin binding domain (SK-FBD) protein coding sequence. The invention relates to a hybrid streptokinase activator (PA) comprises a polypeptide fusion between streptokinase (SK), which are capable of plasminogen (PG) activation, and fibrin binding regions of human fibrinectin, which are from fibrin binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the ability to bind with fibrin independently and also characteristically retains a PG activation ability which becomes evident only after a pronounced duration, or lag, after exposure of the PA to a suitable animal or human PG. The hybrid streptokinase-fibrin binding domain polypeptides are useful in thrombolytic therapy for various kinds of cardiovascular disorders. The hybrids have enhanced fibrin selectivity as well as kinetics of plasminogen activation that are distinct from that of natural streptokinase in being characterised by a temporary delay, or lag of several minutes in the natural rate of the catalytic conversion of plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins can bind tightly with fibrin in blood clots soon after introduction into the vascular system without significantly activating the circulating blood plasminogen to plasmin, thus aiding in the localisation of the plasminogen activation process to the site of pathological thrombus. This overcomes systemic plasminogen activation encountered during clinical use of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4, 08e-176 Length: 2096
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAA37643 (1-2096)

Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
Db 633 AGCCAAATGGTTGGTTCGCTACTGTTGAGGGGACCAATCAAGACATTAGCTTT 692
Qy 21 LysPheGluValAspLeuThrSerArgProAlaHisGlyClyLysThrGluGlnGly 40
Db 693 AAAATTTTGAATCGATTAACATCACGACTGCTCATGGAGGAAAGACAGAGCAAGC 752
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyValaMetSerHisLysLeuGlu 60

Db 753 TTAAGTCCAAAATCAAAACCAATTTGCTACTGATAGTGGCGCGATGTGCACATAAATTGAG 812
Qy 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 813 AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATGATCGCTAACGTCACAGTACGAC 872
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 873 GACTACTTTGAGGTCAATGATTTTGCAGCGCATGCAACCAATTACTATGTCGAAACGGCAAG 932
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 933 GTCTACTTTGCTGACAAAGATGTTCCGTAACCTTGCACCCCAACCTGTCGAAGAAATTT 992
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 993 TTGCTAAGCGGACATGTGCGGTTAGCCATATAAAGAAAAACCAATACAAAACCAAGCG 1052
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 1053 AAATCTGTTGATGCTGGAATATACTGTACAGTTTACTCCCTTAAACCTTGATGACGATTC 1112
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 1113 AGACCAAGGCTCAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGCACCAATCACA 1172
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyLysTyrThr 200
Db 1173 TCTCAAGAATTAAGTCTCAAGCACAAAGCAATTTTAAACAAAACCAACCCAGGCTATACG 1232
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 1233 ATTTATGAACGGTACTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTA 1292
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaIleThrArgLys 240
Db 1293 CCAATCGATCAAGAGTTTACTTTACCGTGTAAAAATCGGGAAACAAGCTTATAGGATCAAT 1352
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlyLysTyrThr 260
Db 1353 AAAAATCTGGTCTGAATGAAGAAATAAACCAACACTGATCTCTGGAATAATATTAC 1412
Qy 261 ValLeuLysLysGlyGlyLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 1413 GTCTTAAAAAAGGGGAAAAGCGGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTT 1472
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 1473 ACCATCAAAATACGTTGATGTCGATACCAACGAATTCGCTAAAAAGTGCAGCAGCTCTTAA 1532
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 1533 GCTAGCGAAACGTAACATTAGACTTCAGAGATTTATACGATCCCTCGTATAGGCTTAACATA 1592
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1593 CTCCTACAAATCTCGATGCTTTTGGTATATAGGACTATACCTTAAGTGAAGTAGAG 1652
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1653 GATAATCAGATGACACCAACCGTATCATAAACCGTTTATATGGGCAAGCGACCCGAAGGA 1712
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1713 GAGAATGCTAGCTACCAATTTAGCT 1736
RESULT 12
ID AAX80497
XX AAX80497 standard; cDNA; 2385 BP.
AC AAX80497;
XX
DT 26-AUG-1999 (first entry)

XX	Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW	streptokinase; fusion protein; ss.
XX	Staphylococcus aureus.
OS	
XX	
XX	
FH	Location/Qualifiers
FT	10. .184
CDS	/*tag= a
FT	/label= FB monomer
FT	185. .358
FT	/*tag= b
FT	/label= FB monomer
FT	359. .1601
FT	/*tag= c
FT	/label= streptokinase
XX	
XX	
XX	US5011686-A.
PN	
XX	
XX	30-APR-1991.
PD	
XX	
XX	15-NOV-1989; 89US-00437769.
XX	
XX	21-SEP-1987; 87US-00099242.
PR	
XX	
XX	(CREA-) CREATIVE BIOMOLEC.
PA	
XX	
XX	Pang RHL; 12
PI	
XX	
XX	WPI; 1991-140198/19.
DR	
DR	P-PSDB; AAR11829.
XX	
XX	Imparting injectable fibrinolytic agent - with affinity for intravascular thrombus, by linking agent to fibrin binding domain.
PT	
PT	Disclosure; Fig 5; 18pp; English.
PS	
XX	
CC	The DNA encodes an FB-FB dimer linked to the streptokinase coding sequence. The FB fragment has selective affinity for fibrin, low affinity for fibrinogen, and minimal immunogenicity, imparting thrombus-targeting capability. See also AAQ11649 and AAQ11650
CC	
CC	
XX	
SQ	Sequence 2030 BP; 673 A; 464 C; 406 G; 487 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	7, 71e-176 Length: 2030
Score:	1898.00 Matches: 367
Percent Similarity:	100.00% Conservative: 1
Best Local Similarity:	Mismatches: 0
Query Match:	Indels: 0
DB:	Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x AAQ11651 (1-2030)	
Qy	1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
:	:
:	:
Db	403 ACCCAATTAGTTGTTAGCGTCTGCTACTGTTAGGGGACGAATCAAGACATTAGTCTT 462
Qy	21 LysPhePheGluIleAspLeuThrSerArgProIleHisGlyGlyysThrGluGlnGly 40
Db	463 AAAATTTTTTGAATTCGATCTAAACATCACGACCTGCTCATGGAGGAAGACAGACGAAGC 522
Qy	41 LeuSerProLysSerLysProPheIleThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db	523 TTAGTCCAAATCAAAACCACTTCTACTGATAGTGGCGCGATGCACTAAACTTGG 582
Qy	61 LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db	583 AAAGCTGACTTACTAAAGGCTATTCAAGAACAAATTGATCGCTAACGTCACAGTAACGAC 642
Qy	81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgHisGlyLys 100
Db	643 GACTACTTTGAGGTCATTGATTTTCAGCGCATGCAACCATTTACTGATCGAAACGGCAAG 702

OS Streptococcus dysgalactiae subsp. equisimilis.
XX Synthetic.

PH Key Location/Qualifiers
XX CDS 1..1242
/*tag= a
/transl_except= (pos:40..42,aa:Asn)
/note= "no stop codon given"

XX US5876999-A.

XX 02-MAR-1999.

XX 06-DEC-1995; 95US-00568393.

XX 06-DEC-1995; 95US-00568393.

XX (NASC-) NAT SCI COUNCIL.

XX Wu H;

XX WPI; 1999-189643/16.

XX P-PSDB; AAX94665.

XX Mutant streptokinase polypeptide - useful as plasmin-resistant
XX thrombolytic agent.

XX Claim 4; Col 11-14; 17pp; English.

XX The present invention describes a mutant streptokinase (SK) polypeptide
XX in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
XX segment of the corresponding native SK is replaced by another amino acid.
XX The present sequence encodes mutant SK K59E. SK is a secretory protein of
XX haemolytic Streptococcus able to activate human plasminogen (HPIg) to
XX plasmin (HPIa), which is a serine protease able to catalyse the
XX hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
XX agent in the treatment of vascular thromboembolytic symptoms such as
XX acute myocardial infarction. Compared with wild-type SK, the K59E mutant
XX is more resistant to degradation by human plasmin and is more effective
XX both in acting as a fibrolytic agent and in activating human plasminogen

SQ Sequence 1242 BP; 421 A; 268 C; 239 G; 314 T; 0 U; 0 Other;

Alignment Scores:

Pred. NO.: 5.22e-176 Length: 1242
Score: 1897.00 Matches: 367
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.73% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x AAX16633 (1-1242)

QY 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspLeuSerLeu 20
DB 46 AGCCAATTAGTTGTTAGGCTTCTCGTACTGTTGAGGGGACCAATCAAGACATTTAGTCTT 105
QY 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
DB 106 AAATTTTGAATCGATCTACATCACCGACCTGCTCATGGAGGAAAGACAGCAAGGC 165
QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
DB 166 TTAAGTCGGGAATCAAAACCACTTTGCTACTGATAGTGGCGCGATGTCACATAAATTTGAG 225
QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
DB 226 AAAGCTGACTTACTAAAGGCTATTCAAGAACCAATTGATCGCTAACTCCACAGTACAGAC 285
QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
DB 286 GACTACTTTGAGTTCATTTGTCAGACGATGCAACCACTTACTGATCGAAGCGCAAG 345

QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
DB 346 GTCTACTTTTGTCTGACAAAGATGTTGCGTAACTTGCCTGACCACTGCTCCAAAGATTTT 405
QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGlyLysProIleGlnAsnGlnAla 140
DB 406 TTGCTAAGCGGACATGTGCGCTTAGACCATATAAAGAAAAACCAATACAAAACCAAGCG 465
QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
DB 466 AAATCTGTTGATGTGAATATATCTGTACAGTTTACTTCCCTTAAACCTGATGACGATTTC 525
QY 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
DB 526 AGACCCAGGTCTCAAGATACTAAGCTATTGAAAAACACTAGCTATCGTGCACCATCACA 585
QY 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
DB 586 TCTCAAGAAATTACTAGCTCAAGCACAAAGCAITTTTAAACAAAAACCAACCCAGCTATACG 645
QY 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
DB 646 ATTATGAACGTGACTCTCTCAATCGTCACTCATGACATGACATTTTCCGTACGATTTTA 705
QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaIleThrIleAsn 240
DB 706 CCATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAACAAGCTTATAGATCAAT 765
QY 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
DB 766 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 825
QY 261 ValLeuLysLysGlyGlyLysProTyrAspProPheAspArgSerHisLysLysLeuPhe 280
DB 826 GTCCCTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTT 885
QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
DB 886 ACATCAATACGTTGATGTCGATACCAACGAATGCTTAAAGATGACGAGCTCTTAAACA 945
QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
DB 946 GCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGAAGGCTAAACTA 1005
QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
DB 1006 CTCTACACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAAAGTAGAG 1065
QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
DB 1066 GATAATCACGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAGGA 1125
QY 361 GluAsnAlaSerTyrHisLeuAla 368
DB 1126 GAGATGCTAGCTATCATTTTAGCC 1149

Search completed: November 6, 2004, 00:37:42
Job time : 616.178 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:33:59 ; Search time 178.959 Seconds
(without alignments)
737.669 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901
Sequence: 1 SQLVSVAGTVEGTNQDISL.....IITVVGKRPGENASYHLA 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23sep04 :
1: Geneseqp1980s :
2: Geneseqp1990s :
3: Geneseqp2000s :
4: Geneseqp2001s :
5: Geneseqp2002s :
6: Geneseqp2003as :
7: Geneseqp2003bs :
8: Geneseqp2004s :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	100.0	413	8	ADL92189 Streptoki
2	1901	100.0	414	2	AAR10194 Streptoki
3	1901	100.0	414	2	AAR63120 Streptoki
4	1901	100.0	414	2	AAY24794 Streptoco
5	1901	100.0	414	2	AAY94664 Streptoco
6	1901	100.0	414	2	AAY01556 Native st
7	1901	100.0	414	3	AAY90282 S. equisi
8	1901	100.0	795	2	AAY24797 Streptoki
9	1898	99.8	531	2	AAR11829 PB-FB-SK
10	1897	99.8	414	2	AAY94665 Streptoco
11	1883.5	99.1	800	2	AAR21723 Plasminog
12	1883.5	99.1	813	2	AAR21728 Wild type
13	1883.5	99.1	1181	2	AAR21727 Streptoki
14	1883.5	99.1	1194	2	AAR21724 Modified
15	1883.5	99.1	1194	2	AAR21726 Streptoki
16	1883.5	99.1	1194	2	AAR21725 Modified
17	1875	98.6	369	2	AAR12892 Truncated
18	1875	98.6	401	3	AAY84004 Amino aci
19	1875	98.6	413	3	AAY84005 Amino aci
20	1875	98.6	414	2	AAR86143 Streptoki
21	1875	98.6	414	3	AAB01295 Wild type
22	1875	98.6	440	2	AAR12889 Streptoki
23	1875	98.6	483	2	AAR12885 Factor xa
24	1875	98.6	483	2	AAR12522 Factor xa
25	1875	98.6	499	2	AAR12891 Streptoki

26	1875	98.6	747	2	AAR12894 Met-core
27	1875	98.6	859	2	AAR12893 OmpAL str
28	1871	98.4	414	2	AAR20202 S.equisim
29	1868.5	98.3	415	3	AAY50870 Streptoco
30	1868.5	98.3	415	3	AAY9593 Streptoco
31	1864	98.1	414	2	AAR86144 De-immuni
32	1864	98.1	414	3	AAB01296 Altered s
33	1859	97.8	440	2	AAR06377 Streptoki
34	1845	97.1	413	2	AAY25020 Streptoki
35	1845	97.1	413	5	ABB80012 Streptoki
36	1845	97.1	413	6	ABG74199 Represent
37	1841	96.8	372	2	AAR10200 Streptoki
38	1833	96.4	372	2	AAR10197 Streptoki
39	1824.5	96.0	371	2	AAR10195 Streptoki
40	1824	95.9	374	2	AAR10198 Streptoki
41	1822	95.8	414	4	AAB74940 Mutant st
42	1819	95.7	384	3	AAY84007 Amino aci
43	1815	95.5	372	3	AAY84006 Amino aci
44	1786	94.0	440	1	AAP50620 Sequence
45	1748.5	92.0	391	2	AAR10199 Streptoki

ALIGNMENTS

RESULT 1
ADL92189
ID ADL92189 standard; protein; 413 AA.
XX
AC ADL92189;
XX
20-MAY-2004 (first entry)
XX
Streptokinase protein sequence.
XX
harvesting; recombinant; host cell; N-terminal leader peptide;
KW pre-peptide; lantibiotic; post-translational modification;
KW pharmaceuticals; vaccine; immunogenic.
XX
Unidentified.
XX
W02003099862-A1.
XX
04-DEC-2003.
XX
26-MAY-2003; 2003WO-NL000389.
XX
24-MAY-2002; 2002EP-00077060.
PR 07-FEB-2003; 2003US-00360101.
XX
(NANO-) APPLIED NANOSYSTEMS BV.
XX
Moll GN, Leenhouts CU, Kuipers OP, Driessen AJM;
WPI; 2004-042770/04.
XX
Harvesting a desired polypeptide produced by a recombinant host cell, for
PT producing pharmaceuticals, comprises selecting a recombinant nucleic acid
PT comprising nucleic acid fragments encoding a leader peptide and the
polypeptide.
XX
Claim 4; Page 82-83; 109pp; English.

The invention relates to a novel method for harvesting a (poly)peptide produced by a recombinant host cell. The novel method involves selecting a cell comprising a first nucleic acid encoding a leader peptide and a second nucleic acid fragment encoding the desired (poly)peptide. The first and second fragments are within the same open reading frame of the first nucleic acid and the leader peptide is functionally equivalent to an N-terminal leader peptide found with the pre-peptide of a lantibiotic. The host cells and nucleic acids are useful for producing, harvesting and post-translational modification of polypeptides. The polypeptides may be used in the production of pharmaceuticals, e.g. as antigen for vaccine or

CC immunogenic composition. This sequence represents a polypeptide relating
CC to the novel method of the invention.

SQ Sequence 413 AA;
Query Match 100.0%; Score 1901; DB 8; Length 413;
Best Local Similarity 100.0%; Pred. No. 4.1e-147;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 75
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQPVQEF 120
Db 76 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQPVQEF 135
Qy 121 LLSGHRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTLLKTLAIGDTIT 180
Db 136 LLSGHRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTLLKTLAIGDTIT 195
Qy 181 SQELLAQAQSIILKNHGGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
Db 196 SQELLAQAQSIILKNHGGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 255
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 256 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db 316 ASERNLDFRDLYDPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375
Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 2
AAR10194
ID AAR10194 standard; protein; 414 AA.

AC AAR10194;

DT 28-MAR-1991 (first entry)

DE Streptokinase encoded by synthetic gene.

KW streptokinase; thrombolytic agent; myocardial infarction.

XX Synthetic.

XX EP407942-A.

XX 16-JAN-1991.

XX 11-JUL-1989; 89JP-00179432.

XX 11-JUL-1989; 89JP-00179432.

XX 27-NOV-1989; 89JP-00307957.

XX 11-APR-1990; 90JP-00096830.

XX (SAKA) OTSUKA PHARM FACTOR.

XX Majima E, Ogino K, Ono K, Sakata Y, Uenoyama T;

XX WPI; 1991-016179/03.

XX N-FSDB; AAQ10230.

XX Synthetic gene encoding streptokinase - scale, high purity prodn. of

XX streptokinase used as a thrombolytic agent.

XX Claim 1; Page 51; 76pp; English.

XX Streptokinase and its derivatives can be produced in large quantities
CC with high purity for use as thrombolytic agents in patients with lung
CC thrombus or myocardial infarction. See also AAR10195-R10200

SQ Sequence 414 AA;

Query Match 100.0%; Score 1901; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 4.1e-147;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 75
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQPVQEF 120
Db 76 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKGGSVTLPTQPVQEF 135
Qy 121 LLSGHRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTLLKTLAIGDTIT 180
Db 136 LLSGHRVRPYKEKPIQNOAKSVDEYTVQFTPLNPDFFRGLKDTLLKTLAIGDTIT 195
Qy 181 SQELLAQAQSIILKNHGGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
Db 196 SQELLAQAQSIILKNHGGTYIYERDSSIVTHNDIFRTILPMDQEFYRVKREQAYRIN 255
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 256 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
Db 316 ASERNLDFRDLYDPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375
Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 3

AAR63120

ID AAR63120 standard; protein; 414 AA.

XX AAR63120;

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 16-NOV-1994 (first entry)

XX Streptokinase.

XX Streptokinase; SK; Streptococcus equisimilis; plasminogen;

KW myocardial infarction.

XX Streptococcus dysgalactiae subsp. equisimilis.

XX Key Location/Qualifiers

FT Region 1. .352

FT Region /note= "claim 3, see CC"

FT Region 14. .414

FT Region /note= "claim 1, see CC"

FT Region 120. .352

FT Region /note= "claim 3, see CC"

FT Region 244. .414

FT Region /note= "claim 3, see CC"

FT Region 244. .352

FT Region /note= "claim 2, see CC"

XX W09407992-A1.

XX PN

XX 14-APR-1994.

XX PD

XX

PF 05-OCT-1993; 93WO-US009502.
 XX
 PR 05-OCT-1992; 92US-00956692.
 PR 29-SEP-1993; 93US-00128299.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Reed GL;
 XX
 DR WPI; 1994-135561/16.
 XX
 PT DNA encoding a polypeptide which binds to plasminogen and corresponds to
 PT region of streptokinase - useful to detect plasminogen in a sample and to
 PT treat myocardial infarction.
 XX
 PS Disclosure; Page 40-41; 62pp; English.
 XX
 CC Nucleic acid comprising a sequence encoding amino acids 14-414 of
 CC streptokinase, which binds to plasminogen and does not have a sequence
 CC comprising amino acids 60-414 is new. The polypeptide pref. comprises
 CC amino acids 244-352, 1-352, 120-352 or 244-414. (Updated on 25-MAR-2003
 CC to correct PN field.) (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 414 AA;
 Query Match 100.0%; Score 1901; DB 2; Length 414;
 Best Local Similarity 100.0%; Pred. No. 4.1e-147;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SOLVSVAGTVEGTNQDISLKFFIEDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 60
 DB 16 SOLVSVAGTVEGTNQDISLKFFIEDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 75
 QY 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 120
 DB 76 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 135
 QY 121 LLSGHRVVRPYKEKPIQNAQKSVDEVYVQFTPLNPDPPGLKDTLLKTLAIGDTIT 180
 DB 136 LLSGHRVVRPYKEKPIQNAQKSVDEVYVQFTPLNPDPPGLKDTLLKTLAIGDTIT 195
 QY 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIERTILPMDQEFYRVKREQAYRIN 240
 DB 196 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIERTILPMDQEFYRVKREQAYRIN 255
 QY 241 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
 DB 256 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
 QY 301 ASERNLDFRDLYPDPRDKAKLLYNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREGE 360
 DB 316 ASERNLDFRDLYPDPRDKAKLLYNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREGE 375
 QY 361 ENASYHLA 368
 DB 376 ENASYHLA 383
 RESULT 4
 ID AAY24794 standard; protein; 414 AA.
 XX
 AC AAY24794;
 XX
 DT 17-OCT-2003 (revised)
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rsk; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;

KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX
 OS Streptococcus dysgalactiae subsp. equisimilis.
 PN WO9931247-A1.
 XX 24-JUN-1999.
 XX
 PD 15-DEC-1998; 98WO-US026694.
 PF 15-DEC-1997; 97US-0069497P.
 PR
 PR (HARD) HARVARD COLLEGE.
 PA
 PI Reed GL;
 XX
 DR WPI; 1999-395183/33.
 DR N-PSDB; AAX80492.
 XX
 PT N-terminally deleted streptokinase.
 XX
 PS Claim 30; Page 60-61; 73pp; English.
 XX
 CC The present invention describes an isolated bacterial protein that
 CC induces fibrin-dependent plasminogen activation in a pharmaceutical
 CC composition for dissolving blood clots. Also described are: (1) a
 CC composition comprising an isolated modified streptokinase, the
 CC modification being removal of amino acid residues in the amino terminus;
 CC (2) a method for dissolving a blood clot in a subject, comprising
 CC administering to the subject a fibrin-dependent streptokinase protein; a
 CC nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
 CC expression vector comprising (1); and (4) a host cell transformed with
 CC the expression vector of (3). The pharmaceutical composition comprising a
 CC bacterial fibrin-dependent plasminogen activator is useful for dissolving
 CC blood clots in patients with a thrombotic condition, e.g. myocardial
 CC infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis,
 CC graft thrombosis and arterial thrombosis. The modified streptokinase can
 CC also be used in non-human mammals. Streptokinase activation of
 CC plasminogen is at least 10-fold, preferably 100-fold greater in the
 CC presence of fibrin than in the absence of fibrin. The modified
 CC streptokinase has at least one amino acid substitution that inactivates a
 CC substrate site for proteolytic cleavage. This reduces the rate of
 CC degradation of the streptokinase at least two-fold. The present sequence
 CC represents native streptokinase (nsk). (Updated on 17-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 414 AA;
 Query Match 100.0%; Score 1901; DB 2; Length 414;
 Best Local Similarity 100.0%; Pred. No. 4.1e-147;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SOLVSVAGTVEGTNQDISLKFFIEDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 60
 DB 16 SOLVSVAGTVEGTNQDISLKFFIEDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 75
 QY 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 120
 DB 76 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 135
 QY 121 LLSGHRVVRPYKEKPIQNAQKSVDEVYVQFTPLNPDPPGLKDTLLKTLAIGDTIT 180
 DB 136 LLSGHRVVRPYKEKPIQNAQKSVDEVYVQFTPLNPDPPGLKDTLLKTLAIGDTIT 195
 QY 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIERTILPMDQEFYRVKREQAYRIN 240
 DB 196 SQELLAQAQSILNKNHPGYTIYERDSSIVTHNDIERTILPMDQEFYRVKREQAYRIN 255
 QY 241 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
 DB 256 KKSGLNEEINNTDLISEKYYVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
 QY 301 ASERNLDFRDLYPDPRDKAKLLYNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREGE 360
 DB 316 ASERNLDFRDLYPDPRDKAKLLYNLDAFGIMDYTLTGKVEDNHDHDTNRIITVYMGKREGE 375
 QY 361 ENASYHLA 368
 DB 376 ENASYHLA 383
 RESULT 4
 ID AAY24794 standard; protein; 414 AA.
 XX
 AC AAY24794;
 XX
 DT 17-OCT-2003 (revised)
 DT 26-AUG-1999 (first entry)
 XX
 DE Streptococcus equisimilis native streptokinase.
 XX
 KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW rsk; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;

Db 316 ASERNLDFRDLPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRITVTVMGRPEG 375
QY 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 5
AAW94664
ID AAW94664 standard; protein; 414 AA.
AC AAW94664;
XX
XX
DT 17-OCT-2003 (revised)
DT 04-MAY-1999 (first entry)
XX
XX
DE Streptococcus equisimilis native streptokinase.
XX
XX
KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.

Key Location/Qualifiers
FH Misc-difference 14
FT /note= "encoded by ACC"
FT
XX
XX
PN US5876999-A.
XX
PD 02-MAR-1999.
XX
XX
PF 06-DEC-1995; 95US-00568393.
XX
PR 06-DEC-1995; 95US-00568393.
XX
XX (NASC-) NAT SCI COUNCIL.
XX
PI Wu H;
XX
XX WPI; 1999-189643/16.
DR N-PsDB; AAX16632.
XX
XX
PT Mutant streptokinase polypeptide - useful as plasmin-resistant
PT thrombolytic agent.
XX
XX Claim 1; Col 7-10; 17pp; English.

The present invention describes a mutant streptokinase (SK) polypeptide in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61 segment of the corresponding native SK is replaced by another amino acid. The present sequence represents native SK. SK is a secretory protein of haemolytic Streptococcus able to activate human plasminogen (HPIg) to plasmin (HPIm), which is a serine protease able to catalyse the hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic agent in the treatment of vascular thromboembolytic symptoms such as acute myocardial infarction. Compared with wild-type SK, the K59E mutant is more resistant to degradation by human plasmin and is more effective both in acting as a fibrinolytic agent and in activating human plasminogen. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 414 AA;
Query Match 100.0%; Score 1901; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 4.1e-147;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVSGTQDISLKFEDLTSRPAHGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVSGTQDISLKFEDLTSRPAHGKTEQGLSPKSPFATDSGAMSHKLE 75

QY 61 KADLLKAIQELIANVHSNDYFEVIDFASDATITDRNGKVPYFADKGSVTLTPQVQEF 120
Db 76 KADLLKAIQELIANVHSNDYFEVIDFASDATITDRNGKVPYFADKGSVTLTPQVQEF 135
QY 121 LLSGHVRVRPYKEKPIQNOAKSVDVYTVQFTPLNPDDEFPRPGLKDTLLKTLAIGDTIT 180
Db 136 LLSGHVRVRPYKEKPIQNOAKSVDVYTVQFTPLNPDDEFPRPGLKDTLLKTLAIGDTIT 195
QY 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDFERTILPMDQEFYRVKNEQAYRIN 240
Db 196 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDFERTILPMDQEFYRVKNEQAYRIN 255
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEOLLT 300
Db 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEOLLT 315
QY 301 ASERNLDFRDLPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRITVTVMGRPEG 360
Db 316 ASERNLDFRDLPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRITVTVMGRPEG 375
QY 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 6
AAW01556
ID AAW01556 standard; peptide; 414 AA.
XX
XX
AC AAW01556;
XX
DT 17-OCT-2003 (revised)
DT 18-JUN-1999 (first entry)
XX
XX
DE Native streptokinase protein sequence.
XX
XX
KW Antigenic peptide; streptokinase; streptokinase-specific antibody;
KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
KW rheumatic fever.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
XX WQ9908698-A1.
XX
XX 25-FEB-1999.
XX
XX 18-AUG-1998; 98WO-US017114.
XX
XX 18-AUG-1997; 97US-0055911P.
XX
XX (HARD) HARVARD COLLEGE.
XX (GEHO) GEN HOSPITAL CORP.
XX
XX Reed GL, Parhami-Seren B;
XX
XX WPI; 1999-190113/16.
XX
XX New polypeptides which bind streptokinase-specific antibodies - useful in
XX thrombolytic therapy.
XX
XX Disclosure; Page 12; 44pp; English.
XX
XX The present sequence represents a native streptokinase. The specification
XX describes a polypeptide which binds to a streptokinase-specific antibody
XX and prevents the antibody binding to native streptokinase. The
XX specification also describes a synthetic polypeptide (PI) comprising an
XX epitope which binds to a streptokinase-specific antibody and reduces
XX thrombolytic activity of streptokinase. PI is used in thrombolytic
XX therapy, and to prevent or treat glomerulonephritis and rheumatic fever.
XX (Updated on 17-OCT-2003 to standardise OS field)
XX
XX Sequence 414 AA;
SQ

Query Match 100.0%; Score 1901; DB 2; Length 414;
 Best Local Similarity 100.0%; Pred. No. 4.1e-147;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVGTTQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSCGAMSHKLE 60
 DB 16 SOLVSVAGTVGTTQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSCGAMSHKLE 75

QY 61 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 120
 DB 76 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 135

QY 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 180
 DB 136 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 195

QY 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
 DB 196 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 255

QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
 DB 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 315

QY 301 ASERNLDFRDLVPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
 DB 316 ASERNLDFRDLVPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375

QY 361 ENASYHLA 368
 DB 376 ENASYHLA 383

RESULT 7

AA190282
 ID AAY90282 standard; protein; 414 AA.
 XX AC AAY90282;
 XX DT 12-SEP-2003 (revised)
 DT 13-OCT-2000 (first entry)
 XX DE S. equisimilis streptokinase.
 XX KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
 KW plasminogen; human; fibrinectin; thrombolytic therapy;
 KW cardiovascular disorder; fibrinectin.
 XX OS Streptococcus dysgalactiae subsp. equisimilis.
 XX PN EPI024192-A2.
 XX PD 02-AUG-2000.
 XX PF 23-DEC-1999; 99EP-00310541.
 XX PR 24-DEC-1998; 98IN-DE003825.
 XX PA (COUL) CSIR COUNCIL SCI IND RES.
 XX PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
 PI Yadav M;
 XX WPI; 2000-516032/47.
 DR N-PSDB; AAA37633.
 XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
 PT thrombolytic therapy comprises a streptokinase fused with fibrin binding
 PT domains of human fibrinectin.
 XX Example 3; Fig 3; 58pp; English.
 PS PS

CC This sequence represents the human Streptococcus equisimilis
 CC streptokinase protein sequence. The invention relates to a hybrid
 CC plasminogen activator (PA) comprises a polypeptide fusion between
 CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
 CC fibrin binding regions of human fibrinectin, which are from fibrin
 CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
 CC ability to bind with fibrin independently and also characteristically
 CC retains a PG activation ability which becomes evident only after a
 CC pronounced duration, or lag, after exposure of the PA to a suitable
 CC animal or human PG. The hybrid streptokinase-fibrin binding domain
 CC polypeptides are useful in thrombolytic therapy for various kinds of
 CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
 CC well as kinetics of plasminogen activation that are distinct from that of
 CC natural streptokinase in being characterised by a temporary delay, or lag
 CC of several minutes in the natural rate of the catalytic conversion of
 CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
 CC can bind tightly with fibrin in blood clots soon after introduction into
 CC the vascular system without significantly activating the circulating
 CC blood plasminogen to plasmin, thus aiding in the localisation of the
 CC plasminogen activation process to the site of pathological thrombus. This
 CC overcomes systemic plasminogen activation encountered during clinical use
 CC of streptokinase. (Updated on 12-SEP-2003 to standardise OS field)
 XX XX
 SQ Sequence 414 AA;

Query Match 100.0%; Score 1901; DB 3; Length 414;
 Best Local Similarity 100.0%; Pred. No. 4.1e-147;
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVGTTQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSCGAMSHKLE 60
 DB 16 SOLVSVAGTVGTTQDISLKPFEDLTSRPAHGGKTEQGLSPKSPKPFATDSCGAMSHKLE 75

QY 61 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 120
 DB 76 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 135

QY 121 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 180
 DB 136 LLSGHVRVRPYKEKPIQNAQSVDEYVQFTPLNPDDEFPRGLKDTLLKTLAIGDTIT 195

QY 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
 DB 196 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 255

QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300
 DB 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 315

QY 301 ASERNLDFRDLVPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 360
 DB 316 ASERNLDFRDLVPRDKAKLLYNNLDFAFGIMDYTLTGKVEDNHDNTNRIITVYMGKRPEG 375

QY 361 ENASYHLA 368
 DB 376 ENASYHLA 383

RESULT 8

AA1924797
 ID AAY24797 standard; protein; 795 AA.
 XX AC AAY24797;
 XX DT 26-AUG-1999 (first entry)
 XX DE Streptokinase and maltose binding protein fusion protein.

XX Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
 KW nSK; rSK; bacterial; blood clot; thrombotic condition;
 KW myocardial infarction; venous thrombosis; pulmonary embolism;
 KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
 XX XX

OS Streptococcus dysgalactiae subsp. equisimilis.
XX Synthetic.

XX WO9931247-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-US026694.

XX 15-DEC-1997; 97US-0069497P.

XX (HARD) HARVARD COLLEGE.

XX Reed GL;

XX WPI; 1999-395183/33.

XX N-PSDB; AAX80497.

XX N-terminally deleted streptokinase.

XX Example; Page 48-51; 73pp; English.

XX The present invention describes an isolated bacterial protein that induces fibrin-dependent plasminogen activation in a pharmaceutical composition for dissolving blood clots. Also described are: (1) a composition comprising an isolated modified streptokinase, the modification being removal of amino acid residues in the amino terminus; (2) a method for dissolving a blood clot in a subject, comprising administering to the subject a fibrin-dependent streptokinase protein; a nucleic acid (I) encoding a modified bacterial streptokinase; (3) an expression vector comprising (1); and (4) a host cell transformed with the expression vector of (3). The pharmaceutical composition comprising a bacterial fibrin-dependent plasminogen activator is useful for dissolving blood clots in patients with a thrombotic condition, e.g. myocardial infarction, venous thrombosis, pulmonary embolism, cerebral thrombosis, graft thrombosis and arterial thrombosis. The modified streptokinase can also be used in non-human mammals. Streptokinase activation of plasminogen is at least 10-fold, preferably 100-fold greater in the presence of fibrin than in the absence of fibrin. The modified streptokinase has at least one amino acid substitution that inactivates a substrate site for proteolytic cleavage. This reduces the rate of degradation of the streptokinase at least two-fold. The present sequence represents a streptokinase and maltose binding protein fusion protein from an example of the present invention

XX Sequence 795 AA;

Query Match 100.0%; Score 1901; DB 2; Length 795;
Best Local Similarity 100.0%; Pred. No. 1e-146;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 60

Db 397 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 456

Qy 61 KADLLKAIQEOQLIANVHSNDDYFEVIDFASDAITDRNGKVYFADKDGSVTLTPQVQEF 120

Db 457 KADLLKAIQEOQLIANVHSNDDYFEVIDFASDAITDRNGKVYFADKDGSVTLTPQVQEF 516

Qy 121 LLSGHVVRVPYKEKPIQNAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 180

Db 517 LLSGHVVRVPYKEKPIQNAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 576

Qy 181 SQELLAQAQSILKNKHPGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVQNRQAYRIN 240

Db 577 SQELLAQAQSILKNKHPGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVQNRQAYRIN 636

Qy 241 KKSGLNEEINNTDLISEKYVILKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 300

Db 637 KKSGLNEEINNTDLISEKYVILKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSQLLT 696

Qy 301 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPEG 360

Db 697 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKRPEG 756

Qy 361 ENASYHLA 368

Db 757 ENASYHLA 764

RESULT 9

AAR11829

ID AAR11829 standard; protein; 531 AA.

XX AAR11829;

XX 08-JUL-1991 (first entry)

XX FB-FB-SK fusion conjugate.

XX Fibrin-binding protein; fibrinolysis; intravascular thrombi; fibrinogen;
KW streptokinase; fusion protein.

XX Staphylococcus aureus.

XX Key Location/Qualifiers
FT Peptide 2..61
FT Peptide /label= PB monomer
FT Peptide 62..119
FT Peptide /label= PB monomer
FT Peptide 120..531
FT Peptide /label= streptokinase

PN US5011686-A.

PD 30-APR-1991.

PF 15-NOV-1989; 89US-00437769.

XX 21-SEP-1987; 87US-00099242.

XX (CREA-) CREATIVE BIOMOLEC.

XX Pang RHL;

XX WPI; 1991-140198/19.

XX N-PSDB; AAQ11651.

XX Imparting injectable fibrinolytic agent - with affinity for intravascular thrombus, by linking agent to fibrin binding domain.

XX Disclosure; Fig 5; 18pp; English.

XX The conjugate comprises an FB-FB dimer linked to streptokinase The FB fragment has selective affinity for fibrin, low affinity for fibrinogen, and minimal immunogenicity, imparting thrombus- targeting capability.
XX See also AAR11821 and AAR11828

XX Sequence 531 AA;

Query Match 99.8%; Score 1898; DB 2; Length 531;
Best Local Similarity 99.7%; Pred. No. 1e-146;
Matches 367; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 60

Db 133 TQLVSVAGTVEGTNQDISLKFEIDLTSRPAHGKTEQGLSPKSPFATDSCAMSHKLE 192

Qy 61 KADLLKAIQEOQLIANVHSNDDYFEVIDFASDAITDRNGKVYFADKDGSVTLTPQVQEF 120

Db 193 KADLLKAIQEOQLIANVHSNDDYFEVIDFASDAITDRNGKVYFADKDGSVTLTPQVQEF 252

Qy 121 LLSGHVVRVPYKEKPIQNAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 180

Db 253 LLSGHVVRVPYKEKPIQNAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 312

QY 181 SOELLAAQASILNKHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
Db 313 SOELLAAQASILNKHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 372
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 373 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 432
QY 301 ASERNLDFRDLYDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNNTNRIITVYMGKPEP 360
Db 433 ASERNLDFRDLYDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNNTNRIITVYMGKPEP 492
QY 361 ENASYHLA 368
Db 493 ENASYHLA 500

RESULT 10
AAW94665
ID AAW94665 standard; protein; 414 AA.
XX
AC AAW94665;
XX
DT 04-MAY-1999 (first entry)
XX
DE Streptococcus equisimilis mutant streptokinase K59E.
XX
KW Streptococcus equisimilis H46A; streptokinase; mutant; fibrinolytic;
KW plasma clot; hydrolysis; haemolytic Streptococcus; plasminogen; plasmin;
KW serine protease; fibrin; blood clot; thrombolytic;
KW vascular thromboembolytic symptom; acute myocardial infarction;
KW fibrinolysis; resistance.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Synthetic.

Key Location/Qualifiers
FH Misc-difference 14
FT /note= "encoded by ACC"
XX
XX US976999-A.
XX
XX 02-MAR-1999.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX 06-DEC-1995; 95US-00568393.
XX
XX (NASC-) NAT SCI COUNCIL.
XX
XX Wu H;
XX
XX WPI; 1999-189643/16.
XX
XX N-PSDB; AAX16633.
XX
XX Mutant streptokinase polypeptide - useful as plasmin-resistant
XX thrombolytic agent.
XX
XX Claim 4; Col 11-14; 17pp; English.
XX
XX The present invention describes a mutant streptokinase (SK) polypeptide
XX in which at least one of the amino acids in the Pro58-Lys59-Ser60-Lys61
XX segment of the corresponding native SK is replaced by another amino acid.
XX The present sequence is mutant SK K59E. SK is a secretory protein of
XX haemolytic Streptococcus able to activate human plasminogen (Hplg) to
XX plasmin (Hplm), which is a serine protease able to catalyse the
XX hydrolysis of fibrin in blood clots. The SK is useful as a thrombolytic
XX agent in the treatment of vascular thromboembolytic symptoms such as
XX acute myocardial infarction. Compared with wild-type SK, the K59E mutant
XX is more resistant to degradation by human plasmin and is more effective
XX both in acting as a fibrinolytic agent and in activating human plasminogen
XX
XX Sequence 414 AA;

Query Match 99.8%; Score 1897; DB 2; Length 414;
Best Local Similarity 99.7%; Pred. No. 8.8e-147;
Matches 367; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SOLVSVAGTVEGTNODISLKFFEIDLTSRPAHGKTEQGLSPKSKPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVEGTNODISLKFFEIDLTSRPAHGKTEQGLSPKSKPFATDSGAMSHKLE 75
QY 61 KADLLKAIQQLANVHSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLPTQPVQEF 120
Db 76 KADLLKAIQQLANVHSNDYFEVIDFASDATITDRNGKYVFADKDGSVTLPTQPVQEF 135
QY 121 LLSGHVVRVPYKEKPIQNAQSVVDVEYTVQFTPLNPDDDFRPGDKTKLKTLAGDTIT 180
Db 136 LLSGHVVRVPYKEKPIQNAQSVVDVEYTVQFTPLNPDDDFRPGDKTKLKTLAGDTIT 195
QY 181 SOELLAAQASILNKHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
Db 196 SOELLAAQASILNKHGPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 255
QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
QY 301 ASERNLDFRDLYDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNNTNRIITVYMGKPEP 360
Db 316 ASERNLDFRDLYDPDRKAKLLYNNLDAFGIMDYTLTGKVEDNHDNNTNRIITVYMGKPEP 375
QY 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 11
AAW21723
ID AAW21723 standard; protein; 800 AA.
XX
AC AAW21723;
XX
DT 17-OCT-2003 (revised)
DT 01-OCT-1997 (first entry)
XX
XX Plasminogen-binding domain from streptokinase.
XX
XX Plasminogen-binding fragment; streptokinase; degradation;
KW thrombolytic agent; blood clot; bolus.
XX
XX Streptococcus dysgalactiae subsp. equisimilis.
XX
XX WO9641883-A1.
XX
XX 27-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009640.
XX
XX 09-JUN-1995; 95US-00488940.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Reed GL;
XX
XX WPI; 1997-065469/06.
XX
XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
XX as thrombolytic agents in treating thrombosis and in medical equipment.
XX
XX Claim 2; Page 35-37; 65pp; English.
XX
XX This sequence represents the plasminogen-binding fragment of
XX streptokinase which lacks the N-terminal 14 amino acids. This modified
XX streptokinase has an in vitro degradation rate at least 2 times slower
XX than that of native streptokinase. Compounds containing modified

CC streptokinases are specifically used as thrombolytic agents for
CC dissolving blood clots in vivo in a mammal, preferably at a dose of 20000
CC U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-
CC OCT-2003 to standardise OS field)

Sequence 800 AA;

Query Match	99.1%;	Score 1883.5;	DB 2;	Length 800;
Best Local Similarity	99.7%;	Pred. No. 2.8e-145;		
Matches 367;	Conservative	0;	Mismatches 0;	Indels 1;

Qy	1	SQLVVSVAGTVEGNTNQDILSLKPFEDILTSRPAHGKTEQGLSPKSPKPPATDSCGAMSHKLE	60
Db	3	SQLVVSVAGTVEGNTNQDILSLKPFEDILTSRPAHGKTEQGLSPKSPKPPATDSCGAMSHKLE	62
Qy	61	KADLLKAIQEOGLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDSVTLLPTQPVQEF	120
Db	63	KADLLKAIQEOGLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDSVTLLPTQPVQEF	122
Qy	121	LLSGHVRVRPYKEKPIQNAQKSDVDVEYTVQFTPLNPDDDFRPLKDTLLKTLAIGDIT	180
Db	123	LLSGHVRVR - YKEKPIQNAQKSDVDVEYTVQFTPLNPDDDFRPLKDTLLKTLAIGDIT	181
Qy	181	SOELLAQAQSLNKNHPGYTIYERDSSIVTHNDIFRTILPMDQBFYVRVKNRBOAYRN	240
Db	182	SOELLAQAQSLNKNHPGYTIYERDSSIVTHNDIFRTILPMDQBFYVRVKNRBOAYRN	241
Qy	241	KKSGLNEEINNTDLISEKYYVLKXGEKYPDPDRSHLKLFTIKYVDVDTNELLKSEQLLT	300
Db	242	KKSGLNEEINNTDLISEKYYVLKXGEKYPDPDRSHLKLFTIKYVDVDTNELLKSEQLLT	301
Qy	301	ASERNLDFRDLYDPRDKAKLYNNLDAGFIMDYTLTGKVEDNHDDDTNRIITYVMGKRPEG	360
Db	302	ASERNLDFRDLYDPRDKAKLYNNLDAGFIMDYTLTGKVEDNHDDDTNRIITYVMGKRPEG	361
Qy	361	ENASYHLA 368	
Db	362	ENASYHLA 369	

RESULT 12

AAW21728
ID AAW21728 standard; protein; 813 AA.

XX AAW21728:

DT 17-QCT-2003 (revised)

DI	17-OCT-2003	(revised)
DT	01-OCT-1997	(first entry)

DE Wild type plasminogen-binding fragment of Streptokinase.

XX Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW thrombolytic agent; blood clot; bolus; maltose-binding protein.
KW thrombolytic agent; blood clot; bolus; maltose-binding protein.

XX. Streptococcus dysgalactiae subsp. equisimilis.

XX PN WO9641883-A1.

XX
PD 27-DEC-1996.

07-JUN-1996; 96WO-US009640.

XX
PR 09-JUN-1995; 95US-00488940.

PA (HARD) HARVARD COLLEGE.

PI Reed GL;

WPI; 1997-065469/06.

Modified forms of streptokinase resistant to enzymatic cleavage - useful as thrombolytic agents in treating thrombosis and in medical equipment.

Example 1; Page 12-13; 65pp; English.

This sequence represents the wild type plasminogen-binding fragment of streptokinase. This fragment was used in the design of a modified streptokinase has an *in vitro* degradation rate at least 2 times slower than that of native streptokinase. Compounds containing modified streptokinases are specifically used as thrombolytic agents for dissolving blood clots *in vivo* in a mammal, preferably at a dose of 20000 U/kg, opt. as a bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 813 AA;

Query Match 99.1%; Score 1883.5; DB 2; Length 813;
Best Local Similarity 99.7%; Pred. No. 2.9e-145;
Matches 367; Conservative 0; Mismatches 0; Indels 1;

Qy	1	SOLVSVAGTVEGTNQDII	SLKPF	EIDLTSRPAHGKTQEGLSPKPPATDSGAMSHKLE	60
Db	16	SOLVSVAGTVEGTNQDII	SLAF	FEDLTSRPAHGKTQEGLSPKPPATDSGAMSHKLE	75
Qy	61	KADILKAIQEOQLIANVHSNDY	FEVIDFAS	DATIITDRNGKVVFADKGVSVTLPQVBQEF	120
Db	76	KADILKAIQEOQLIANVHSNDY	FEVIDFAS	DATIITDRNGKVVFADKGVSVTLPQVBQEF	135
Qy	121	LLSGHVRVPYKEKPIQNOAKSVDVEYTVQFTPLNPDDDFRGLKDXTLLKTLAIGDTTIT	180		
Db	136	LLSGHVRV-YKEKPIQNOAKSVDVEYTVQFTPLNPDDDFRGLKDXTLLKTLAIGDTTIT	194		
Qy	181	SOELLAQAQSILKNKHPGYTIYERDSSI	VTHDNDIFRTILPMQDBFTYVQNRQEQAYRN	240	
Db	195	SOELLAQAQSILKNKHPGYTIYERDSSI	VTHDNDIFRTILPMQDBFTYVQNRQEQAYRN	254	
Qy	241	KKSGLNEEINNTDLISEKYVYLKXGEKPYDPDRSHLKLFTTKYDVVDNTELLKSQLLT	300		
Db	255	KKSGLNEEINNTDLISEKYVYLKXGEKPYDPDRSHLKLFTTKYDVVDNTELLKSQLLT	314		
Qy	301	ASERNLOFRDLYDPDRDAKALYNLNLD	AFGIMDYTLTGKVEDNHDDTNRIITYVMGRKRPEG	360	
Db	315	ASERNLOFRDLYDPDRDAKALYNLNLD	AFGIMDYTLTGKVEDNHDDTNRIITYVMGRKRPEG	374	
Qy	361	ENASYHLA	368		
Db	375	ENASYHLA	382		

RESULT 13

AAW21727

ID AAW21727 standard; protein; 1181 AA.

AAW21727;

17-OCT-2003 (revised)

DT	01-OCT-1997	(first entry)
DI	17-OCT-2003	(revised)

XX	Streptokinase/maltose binding protein fusion protein, rSKdelta14;
DE	
XX	
KW	Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW	thrombolytic agent; blood clot; bolus; maltose-binding protein.

XX
QS *Streptococcus dysgalactiae* subsp. *equisimilis*.

XX	Key	Location/Qualifiers
FH		

FT	key	LOCATION
FT	Protein	1. .381

FT /label= Maltose_binding_protein

FT /note= "acts as blocking group"

FT	Protein	382.1
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

FT	label= Modified_Streptokinase
rm	/rest= "res N-terminal 14 amino acids deleted"

FT yy /note= "Has N-terminal 14 amino acids deleted"

XX PN WO9641883-A1

XX
PD
27-DEC-1996

XX 07-JUN-1996; 96WO-US009640.
XX 09-JUN-1995; 95US-00488940.
XX (HARD) HARVARD COLLEGE.
XX
XX Reed GL;
XX WPI; 1997-065469/06.
XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
XX as thrombolytic agents in treating thrombosis and in medical equipment.
XX
XX Example 1; Page 12; 65pp; English.
XX
XX This sequence represents a fusion protein between maltose-binding protein
XX and a mutant form of the plasminogen-binding fragment of
XX streptokinase which has the N-terminal 14 amino acids deleted. This
XX fusion protein was used in the design of a modified streptokinase has an
XX in vitro degradation rate at least 2 times slower than that of native
XX streptokinase. Compounds containing modified streptokinases are
XX specifically used as thrombolytic agents for dissolving blood clots in
XX vivo in a mammal, preferably at a dose of 20000 u/kg, optionally as a
XX standardise OS field)
XX
XX Sequence 1181 AA;
XX
XX Query Match 99.1%; Score 1883.5; DB 2; Length 1181;
XX Best Local Similarity 99.7%; Pred. No. 4.9e-145;
XX Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
XX 1 SOLVSVAGTVEGTNODISLKFEIDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 60
XX 384 SOLVSVAGTVEGTNODISLKFEIDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 443
XX
XX 61 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLTPQVQEF 120
XX 444 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLTPQVQEF 503
XX
XX 121 LLSGHVRVRPYKEKPIQNAQSVVDVEYVQFTPLNPDPPRGLKDTLKTLLAIGDTIT 180
XX 504 LLSGHVRVR-YKEKPIQNAQSVVDVEYVQFTPLNPDPPRGLKDTLKTLLAIGDTIT 562
XX
XX 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 240
XX 563 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 622
XX
XX 241 KKSGLNEEINNTDLISEKYVLLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
XX 623 KKSGLNEEINNTDLISEKYVLLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 682
XX
XX 301 ASERNLDFRDLYDPRDKAKLLYNLLDAFGIMDYTLTGKVEDNHDNRIITVYMGKREPEG 360
XX 683 ASERNLDFRDLYDPRDKAKLLYNLLDAFGIMDYTLTGKVEDNHDNRIITVYMGKREPEG 742
XX
XX 361 ENASTHLA 368
XX 743 ENASTHLA 750
XX
XX RESULT 14
XX AAW21724
XX ID AAW21724 standard; protein; 1194 AA.
XX
XX AC AAW21724;
XX
XX 01-OCT-1997 (first entry)
XX
XX Modified streptokinase, rsk5mut.
XX
XX Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW

thrombolytic agent; blood clot; bolus; maltose-binding protein.
OS Streptococcus dysgalactiae subsp. equisimilis.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Protein 1..381
XX /label= Maltose binding protein
XX /note= "acts as blocking group"
XX Protein 382..1194
XX /label= Modified streptokinase
XX Misc-difference 391
XX /label= R10A
XX Misc-difference 417
XX /label= R36A
XX Misc-difference 426
XX /label= R45A
XX Misc-difference 432
XX /label= R51A
XX Misc-difference 440
XX /label= R59A
XX
XX WO9641883-A1.
XX
XX 27-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US009640.
XX
XX 09-JUN-1995; 95US-00488940.
XX (HARD) HARVARD COLLEGE.
XX
XX Reed GL;
XX WPI; 1997-065469/06.
XX
XX Modified forms of streptokinase resistant to enzymatic cleavage - useful
XX as thrombolytic agents in treating thrombosis and in medical equipment.
XX
XX Claim 15; Page 41-44; 65pp; English.
XX
XX This sequence represents a fusion protein between maltose-binding protein
XX and a modified form of the plasminogen-binding fragment of streptokinase
XX containing 5 point mutations. This modified streptokinase has an in vitro
XX degradation rate at least 2 times slower than that of native
XX streptokinase. Compounds containing modified streptokinases are
XX specifically used as thrombolytic agents for dissolving blood clots in
XX vivo in a mammal, preferably at a dose of 20000 u/kg, optionally as a
XX bolus rather than by continuous infusion
XX
XX Sequence 1194 AA;
XX
XX Query Match 99.1%; Score 1883.5; DB 2; Length 1194;
XX Best Local Similarity 99.7%; Pred. No. 4.9e-145;
XX Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
XX
XX 1 SOLVSVAGTVEGTNODISLKFEIDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 60
XX 797 SOLVSVAGTVEGTNODISLKFEIDLTSRPAHGKTEQGLSPKSPFATDGSAMSHKLE 856
XX
XX 61 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLTPQVQEF 120
XX 857 KADLLKAIQELIANVHNSDDYFEVIDFASDATITDRNGKVYFADKDGSVTLTPQVQEF 916
XX
XX 121 LLSGHVRVRPYKEKPIQNAQSVVDVEYVQFTPLNPDPPRGLKDTLKTLLAIGDTIT 180
XX 917 LLSGHVRVR-YKEKPIQNAQSVVDVEYVQFTPLNPDPPRGLKDTLKTLLAIGDTIT 975
XX
XX 181 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 240
XX 976 SQELLAQAQSILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 1035
XX
XX 241 KKSGLNEEINNTDLISEKYVLLKKGKPYDPPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300

```
Db 1036 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 1095
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKRPEG 360
Db 1096 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKRPEG 1155
Qy 361 ENASYHLA 368
Db 1156 ENASYHLA 1163

RESULT 15
AAW21726
ID AAW21726 standard; protein; 1194 AA.
XX
AC AAW21726;
XX
DT 17-OCT-2003 (revised)
DT 01-OCT-1997 (first entry)
XX
DE Streptokinase/maltose binding protein fusion protein, rSK.
XX
KW Plasminogen-binding fragment; streptokinase; degradation; MBP;
KW thrombolytic agent; blood clot; bolus; maltose-binding protein.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
XX
FH Key Location/Qualifiers
FT Protein 1..381
FT /label= Maltose binding protein
FT /note= "acts as blocking group"
FT 382..1194
FT /label= Streptokinase
XX
FN WO9641883-A1.
XX
PD 27-DEC-1996.
XX
PF 07-JUN-1996; 96WO-US009640.
XX
PR 09-JUN-1995; 95US-00488940.
XX
PA (HARD ) HARVARD COLLEGE.
XX
PI Reed GL;
XX
WPI; 1997-065469/06.
XX
PT Modified forms of streptokinase resistant to enzymatic cleavage - useful
PT as thrombolytic agents in treating thrombosis and in medical equipment.
XX
PS Example 1; Page 11-12; 65pp; English.
XX
CC This sequence represents a fusion protein between maltose-binding protein
CC and the plasminogen-binding fragment of streptokinase. This fusion
CC protein was used in the design of modified streptokinase has an in vitro
CC degradation rate at least 2 times slower than that of native
CC streptokinase. Compounds containing modified streptokinases are
CC specifically used as thrombolytic agents for dissolving blood clots in
CC vivo in a mammal, preferably at a dose of 20000 U/kg, optionally as a
CC bolus rather than by continuous infusion. (Updated on 17-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 1194 AA;
```

```
Query Match 99.1%; Score 1883.5; DB 2; Length 1194;
Best Local Similarity 99.7%; Pred. No. 4.9e-145;
Matches 367; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 SOLVSVAGTVGTTNODISLKFFEDLTSRPAHGKTEQGLSPKSPFATDSCGAMSHKLE 60
Db 397 SOLVSVAGTVGTTNODISLKFFEDLTSRPAHGKTEQGLSPKSPFATDSCGAMSHKLE 456
```

```
Qy 61 KADLLKAIQEOLIANVHSNDDYFEVIDFASDATITDRNGKVYPADKDGSVTLTQPQVEF 120
Db 457 KADLLKAIQEOLIANVHSNDDYFEVIDFASDATITDRNGKVYPADKDGSVTLTQPQVEF 516
Qy 121 LLSGHRVVRPYKEKPIQONAKSVVDVEYTVQFTPLNPDDEFPRPGLKDTKLLKTLAIGDTIT 180
Db 517 LLSGHRVVR - YKEKPIQONAKSVVDVEYTVQFTPLNPDDEFPRPGLKDTKLLKTLAIGDTIT 575
Qy 181 SQELLAQAQSILNKNHPGYTIYERDSSIYTHDNDIERTILPMDOEFTYRVKNREQAYRIN 240
Db 576 SQELLAQAQSILNKNHPGYTIYERDSSIYTHDNDIERTILPMDOEFTYRVKNREQAYRIN 635
Qy 241 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 636 KKSGLNEEINNTDLISEKYVVLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 695
Qy 301 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKRPEG 360
Db 696 ASERNLDFRDLYDPRDKAKLLYNNLDAFGIMDYTLTGKVEDNHDNRIITVYMGKRPEG 755
Qy 361 ENASYHLA 368
Db 756 ENASYHLA 763
```

Search completed: November 3, 2004, 23:49:22
Job time : 181.959 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2004, 02:55:27 ; Search time 632.658 Seconds
(without alignments)
3131.797 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901

Sequence: 1 SQLVSVAGTVEGTNQDISL.....IITVNGKRPGENASYHLA 368

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_03112004_174041_11368/app_query.fasta_1.1045
-DB=PublishedApplications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USERS=US09940235 @CGN 1.1 1034 @runat_03112004_174041_11368
-NCPU=6 -ICPU=3 -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
20: /cgn2_6/ptodata/2/pubpna/US6_NEW_PUB.seq:
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB ID	Description
1	1901	100.0	1245	10	US-09-940-235-1
2	1901	100.0	1327	10	US-09-940-235-6
3	1901	100.0	1377	10	US-09-940-235-5
4	1901	100.0	1541	10	US-09-940-235-9
5	1901	100.0	1661	10	US-09-940-235-10
6	1901	100.0	1782	10	US-09-940-235-11
7	1901	100.0	2096	10	US-09-940-235-12
8	140.5	7.4	7155	16	US-10-335-977-3325
9	138.5	7.3	7320	16	US-10-335-977-3326
10	118.5	6.2	3714	16	US-10-398-221-2298
11	118.5	6.2	3735	16	US-10-398-221-681
12	118.5	6.2	1163020	16	US-10-398-221-10
13	118.5	6.2	3011208	16	US-10-398-221-2058
14	118	6.2	1323	16	US-10-282-122A-8853
15	117	6.2	2332	16	US-10-425-114-3480
16	117	6.2	2655	16	US-10-425-114-35750
17	117	6.2	3129	18	US-10-425-115-59262
18	115.5	6.1	16680	16	US-10-282-122A-38993
19	113.5	6.0	1683	16	US-10-282-122A-35391
20	111.5	5.9	14067	16	US-10-282-122A-40681
21	110.5	5.8	4953	17	US-10-437-963-23967
22	110	5.8	2406	16	US-10-257-174-5
23	110	5.8	3017	16	US-10-302-172-153
24	110	5.8	3398	16	US-10-424-599-97505
25	109.5	5.8	3552	16	US-10-282-122A-17070
26	109.5	5.8	150223	17	US-10-624-149A-1
27	109	5.7	2499	16	US-10-398-221-1067
28	109	5.7	2499	16	US-10-398-221-2637
29	109	5.7	3897	14	US-10-082-830-152
30	109	5.7	6042	16	US-10-398-221-2044
31	109	5.7	6042	16	US-10-282-122A-24424
32	109	5.7	3673778	15	US-10-312-841-2
33	108.5	5.7	951	16	US-10-398-221-3072
34	108.5	5.7	3326	8	US-08-781-986A-442
35	108.5	5.7	3326	16	US-10-329-624-442
36	108	5.7	1077	9	US-09-974-300-2859
37	108	5.7	2916	16	US-10-282-122A-17169
38	108	5.7	3827	9	US-09-870-759-44
39	108	5.7	3827	10	US-09-751-708A-44
40	108	5.7	3827	18	US-10-428-817A-40
41	108	5.7	9899	15	US-10-311-455-2415
42	107.5	5.7	2499	16	US-10-282-122A-16612
43	107	5.6	1518	16	US-10-282-122A-36130
44	107	5.6	3492	16	US-10-282-122A-15680
45	106.5	5.6	2823	16	US-10-398-221-2045

ALIGNMENTS

RESULT 1

US-09-940-235-1
; Sequence 1, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23

```
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1245
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1242)
US-09-940-235-1
Alignment Scores:
Pred. No.: 2,828-179 Length: 1245
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x US-09-940-235-1 (1-1245)
Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 46 AGCCAATTAGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGNAATCAAGACATTAGTCTT 105
Qy 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 106 AAATTTTGAATCGATCTAATCATCAGCACCTGCTCATGGAGGAAAGACAGAGCAAGGC 165
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 166 TTAAGTCCAAATCAAAACCAATTTGCTACTGATGTAGTGGCGGCGATGTCACATAAATGAG 225
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGlnGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 226 AAAGCTGACTTACTAAGAGCTATTCAAGACAAATGATGCTTAACGTCACAGTAACGAC 285
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 286 GACTACTTTGAGTCAATGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGGCAAG 345
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 346 GTCTACTTTGCTGACAAAGATGGTTGCGTAACCTTGCAGACCAACCTGCTCCAAGAAATTT 405
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 406 TTGCTAACGGACATGTCGGCGTTAGACCATATTAAGAAAAAACCAATACAAAACCAAGCG 465
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 466 AAATCTGTTGATGGGAATATACTGTACAGTTTACTCCCTTAAACCCCTGATGACGATTC 525
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 526 AGACCAGGTCTCAAGAGACTAAGCTATTGAAACACTAGCTATCGTGACACCAATCACA 585
Qy 181 SerGlnGlnLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 586 TCTCAAGAAATTTACTAGCTCAAGCAAAAGCAATTTTAAACAAAACCCAGCTATACG 645
Qy 201 IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 646 ATTTATGAACGTGACTCTCAATCGTCACTCATGACAAATGACATTTTCCGTACGATTTTA 705
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 706 CCAATGGATCAAGAGTTTACTTACCGTGTATTAATAATCGGGAACCAAGCTTATAGGATCAAT 765
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 766 AAAAAATCTGGTCTGAATGAAGAAATAAACACACACTGACCTGATCTCTGAGAAATATTAC 825
Qy 261 ValLeuLysLysGlyGluLysProTyrAspPheAspArgSerHisLeuLysLeuPhe 280
Db 826 GTCCTTAAAAAAGGGGAAGCCGATGATGATCCCTTTGATCGCAGTCACITTGAAACCTGTT 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAATACCTGTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACA 945
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTAGCGAAACGTAACTTAGACTTCAGAGATTTATACGATCTCTCGTATAGGCTAAACTA 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTCCTACAAACATCTCGATGCTTTTGTGATTTATGAGTACTATACCTTAACGTGAAAAGTAGAG 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGA 1125
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGAATGCTAGCTATCATTTAGCC 1149
RESULT 2
US-09-940-235-6
; Sequence 6, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Streptococcus equisimilis
US-09-940-235-6
Alignment Scores:
Pred. No.: 3,088-179 Length: 1327
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-940-235-2_COPY_16_383 (1-368) x US-09-940-235-6 (1-1327)
Qy 1 SerGlnLeuValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 128 AGCCAATTAGTTGTTAGCGTTGCTGCTACTGTTGAGGGGACGNAATCAAGACATTAGTCTT 187
Qy 21 LysPheGluIleAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 211 AAAAAATCTGGTCTGAATGAAGAAATAAACACACACTGACCTGATCTCTGAGAAATATTAC 825
Qy 261 ValLeuLysLysGlyGluLysProTyrAspPheAspArgSerHisLeuLysLeuPhe 280
Db 826 GTCCTTAAAAAAGGGGAAGCCGATGATGATCCCTTTGATCGCAGTCACITTGAAACCTGTT 885
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 886 ACCATCAATACCTGTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAACA 945
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 946 GCTAGCGAAACGTAACTTAGACTTCAGAGATTTATACGATCTCTCGTATAGGCTAAACTA 1005
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1006 CTCCTACAAACATCTCGATGCTTTTGTGATTTATGAGTACTATACCTTAACGTGAAAAGTAGAG 1065
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1066 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCGACCCGAAAGGA 1125
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1126 GAGAATGCTAGCTATCATTTAGCC 1149
```



```
Qy 181 SerGlnGluLeuLeuAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 718 TCTCAAGAACTACTAGCTCAAGCAACAGCAATTTAAACAAACACCACCCAGGCTATACG 777
Qy 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 778 ATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTA 837
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 838 CCAATGGATCAAGAGTTTACTTACCGTGTATAAAATCGGGAACAAGCTTATAGGATCAAT 897
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 898 AAAAAATCTGCTGAATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTAC 957
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 958 GTCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACCTTGAACCTGTTTC 1017
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 1018 ACCATCAAAATCGTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAAACA 1077
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 1078 GCTAGCGAAGTAACTTAGACTTCAGAGATTTATAGCATCTTCGGTAGAAGCTAAACTA 1137
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1138 CTCTACAACAACTCGATGCTTTGGTATTATGGACTATACCTTAACTGGAAGATAGAG 1197
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1198 GATAATCAGCATGACACCAACCGTATCATACCGTTTATATGGCAAGCGACCCGAAGGA 1257
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1258 GAGAAATGCTAGTATCATTTAGGCC 1281

RESULT 4
US-09-940-235-9
; Sequence 9, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1541
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
; US-09-940-235-9
```

```
Alignment Scores:
Pred. No.: 3,77e-179 Length: 1541
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-09-940-235-9 (1-1541)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 78 AGCCAAATGGTTGTAGCGTTGCTGCTACTGTTGAGGGAGCAATCAAGACATTAAGTCTT 137
Qy 21 LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
Db 138 AAAATTTTTTGAATCGATCTAACATCACGACTGCTCATGGAGGAAAGACAGACGAAGGC 197
Qy 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 198 TTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGTCACATAAACTTGAG 257
Qy 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
Db 258 AAAGCTGACTTACTTAAGGCTATTCAAGAAACAATTGATCGCTAAACGTCACAGTAACGAC 317
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 318 GACTACTTTTGAGGTCAATTGATTTTGCAGCGATGCAACCATTTACTGATCGAAACCGCAAG 377
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 378 GTCTACTTTGCTGCAAGATGTTGCGTAACTTGGCCGACCAACCTGTCCAGGAAATTT 437
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 438 TTGCTAAGCGGACATGTCGCGCTTAGACCATATAAGAAAAAACCAATACAAACCAAGCG 497
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 498 AAATCTGTTGATGTGAATATATCTGTACAGTTTACTTCCCTTAAACCCCTGATCGGATTC 557
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 558 AGACCAAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACCATCACA 617
Qy 181 SerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 618 TCTCAAGAACTACTAGCTCAAGCAACAAAGCAATTTTAAACAAACCAACCCAGGCTATACG 677
Qy 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 678 ATTTATGAACGTGACTCTCTCAATCGTCACTCATGACAAATGACATTTTCGGTACGATTTTA 737
Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 738 CCAATGGATCAAGAGTTTACTTACCGTGTATAAAATCGGGAACAAGCTTATAGGATCAAT 797
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 798 AAAAAATCTGCTGCTGATGAAGAAATAACAACACTGACCTGATCTCTGAGAAATATTAC 857
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe 280
Db 858 GTCTTAAAAAAGGGGAAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTTTC 917
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeuThr 300
Db 918 ACCATCAAAATCGTTGATGTCGATACCAACGAATTGCTAAAAAGTGAGCAGCTCTTAAACA 977
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
```


; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-11

Alignment Scores:
Pred. No.: 4,598-179 Length: 1782
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-09-940-235-11 (1-1782)

Qy 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
Db 583 AGCCAAATTTGGTTAGCGTTCTGCTGTTACTGTTGGGACCAATCAAGACATTAAGTCTT 642
Qy 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyThrGluGlnGly 40
Db 643 AAAATTTTGGAAATCGATTAACATCAGCACCTGCTCATGGAGGAAAGACAGCAAGGC 702
Qy 41 LeuSerProLysSerIlyProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
Db 703 TTAAGTCCAAATCAAAACCATTTGCTACTGATAGTGGCGCATGTGCATATAAATTTAG 762
Qy 61 LysAlaAspLeuLeuLysAlaIleGlnGlnGlnLeuLeuAlaAsnValHisSerAsnAsp 80
Db 763 AAAGCTGACTTACTAAGAGCTATTCAAGAACATTTGATCGTAACTCCACAGTACAGAC 822
Qy 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
Db 823 GACTACTTTGAGTCAATGATTTTGAAGCGATGCAACCATTAATGATCGATAACGCAAG 882
Qy 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
Db 883 GTCTACTTTGCTGACAAAGATGGTTGCGTAACCTTGCCGACCAACCTGTCCAAAGATTT 942
Qy 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
Db 943 TTGCTAAGCGGACATGTGGCGCTTAGACCATATAAGAAAAACCAATACAAACCAAGCG 1002
Qy 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
Db 1003 AAATCTGTTGATGCGAATATATCTGTACAGTTTACTCCCTTAAACCTGTACGCGATTTC 1062
Qy 161 ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr 180
Db 1063 AGACCAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCA 1122
Qy 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
Db 1123 TCTCAAGAAATTAAGTACGACCAAGCAATTTTAAACAAACCAACCCAGCGCTATACG 1182
Qy 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
Db 1183 ATTTATGAACGTGATCTCTCATCTCACTCATGACATGACATTTTCCGTACGATTTTA 1242

Qy 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
Db 1243 CCAATGGATCAAGAGTTTACTTACCGTGTAAATAATCGGGAACAACCTTATAGATCAAT 1302
Qy 241 LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr 260
Db 1303 AAAAAATCTGGTCTGAATGAAGAAATAAACACACACTGACCTGATCTCTGAGAAATATTAC 1362
Qy 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLysLysLeuPhe 280
Db 1363 GTCCTTTAAAAAAGGGGAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAACCTGTT 1422
Qy 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
Db 1423 ACCATCAATACATGTTGATGTGCATACCAACAGATTTGCTAAAGATGAGCAGCTTTAA 1482
Qy 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
Db 1483 GCTAGCGAACGTAACTTACGATTTACAGATTTTATACGATCTCGTGAAGGCTAAACTA 1542
Qy 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
Db 1543 CTCTACACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGAAAAAGTAGAG 1602
Qy 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360
Db 1603 GATAATCAGATGACACCAACCGTATCATACCGTTTATATGGGCAAGCAGCCGCAAGGA 1662
Qy 361 GluAsnAlaSerTyrHisLeuAla 368
Db 1663 GAGAATGCTAGCTATCATTTAGCC 1686

RESULT 7

US-09-940-235-12
; Sequence 12, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2096
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-12

Alignment Scores:
Pred. No.: 5,738-179 Length: 2096
Score: 1901.00 Matches: 368
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-2_COPY_16_383 (1-368) x US-09-940-235-12 (1-2096)

QY 1 SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu 20
DB 633 AGCCAAATTGGTTGTTAGCGTTCGTACTGTTGAGGGACGAATCAAGACATTAGCTCT 692

QY 21 LysPhePheGluLeuAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly 40
DB 693 AAAATTTTGAATCGATTAACATCACACGCTGCTCATGGAGGAAAGACAGACGAAGC 752

QY 41 LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGlu 60
DB 753 TTAAGTCAAAATCAAAACCATTTGCTACTGATGAGGCGGCGATGCACATAAATGAG 812

QY 61 LysAlaAspLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp 80
DB 813 AAAGCTGACTTAAGAGCTATTCAAGAACCAATTGATCGCTAACGTCACAGTACAGAC 872

QY 81 AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys 100
DB 873 GACTACTTTGAGGTCAATGATTTTGCAGCGATGCAACCATTTACTGATCGAAACGCAAG 932

QY 101 ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe 120
DB 933 GTCTACTTTGCTGACAAAGATGGTTGCGTAACCTTGCCGACCAACCTGTCCAAGAAATTT 992

QY 121 LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla 140
DB 993 TTGCTAAGCGGACATGTCGCGGTAGACCATATAAGAAACCAATACAAACCAAGCG 1052

QY 141 LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspPhe 160
DB 1053 AAATCTGTTGATGGATATACTGTACAGTTTACTCCCTTAAACCTTGATGACGATTTC 1112

QY 161 ArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIleGlyAspThrIleThr 180
DB 1113 AGACAGGTCTCAAGATACTAAGCTATTGAAACACTAGCTATCGGTGACACCATCACA 1172

QY 181 SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThr 200
DB 1173 TCTCAAGAAATTAAGTCTAGCTCAAGCAAGCATTTTAAACAAACCAACCGAGCTATACG 1232

QY 201 IleTyrGluArgAspSerIleValThrHisAspAsnAspIlePheArgThrIleLeu 220
DB 1233 ATTTATGAACGTGACTCTCATCTGCTCAATGATGACATGACATTTTCGTCAGATTTTA 1292

QY 221 ProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsn 240
DB 1293 CCAATGGATCAAGAGTTTACTTACCGTGTAAAAATCGGGAACAAGCTTATAGGATCAAT 1352

QY 241 LysLysSerGlyLeuAsnGluLeuIleAsnAsnThrAspLeuIleSerGluLysTyrThr 260
DB 1353 AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC 1412

QY 261 ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLysLysLeuPhe 280
DB 1413 GTCTTTAAAAAGGGAAAGCCGATGATCCTTTGATCGGAGTCACCTTGAACATGTTTC 1472

QY 281 ThrIleLysTyrValAspValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr 300
DB 1473 ACCATCAATAGTTGATGTCATCAACCAAGATTGCTTAAAGTGAGCAGCTCTTAAACA 1532

QY 301 AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu 320
DB 1533 GCTAGCGAACGTAATTAGACTTCAGAGATTTATACGATCTCGTGTAAAGGCTAAACTA 1592

QY 321 LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu 340
DB 1593 CTCTACACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACCTGGAAGTAGAG 1652

QY 341 AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly 360

DB 1653 GATAATCACGATGACACCAACCGTATCATACCGTTTATATGGCAAGCCGCAAGGA 1712

QY 361 GluAsnAlaSerTyrHisLeuAla 368
DB 1713 GAGAATGCTAGCTACCATTTAGCT 1736

RESULT 8
US-10-335-977-3325
; Sequence 3325, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 3325:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...7155
; SEQUENCE DESCRIPTION: SEQ ID NO: 3325:
US-10-335-977-3325

Alignment Scores:
Pred. No.: 0.00345 Length: 7155
Score: 140.50 Matches: 82
Percent Similarity: 36.93% Conservative: 65
Best Local Similarity: 20.60% Mismatches: 128
Query Match: 7.39% Indels: 123
DB: 16 Gaps: 16

US-09-940-235-2_COPY_16_383 (1-368) x US-10-335-977-3325 (1-7155)

QY 21 LysPhePheGluLeuAspLeuThr-----SerArg 30
DB 3046 AAAGACTTTGAAAAAGAAATAACCCCTAGAAAGCGCTTAAATGCAAAACATAGAACCAAGG 3105

QY 31 ProAlaHisGlyGlyLysThr-----GluGlnGlyLeuSerProLysSerLys 46

```
Db 3106 CAAGCTCAAGCCAAAGAGCTCAATCTTTTGAAGAGCTTTAAACCTTAAAGAA 3165
Qy 47 PropheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLys 66
Db 3166 CTTATTATCACTAAGCCT-----AAAGAGCATTAATTGCA 3201
Qy 67 AlaIleGlnGluLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIle 86
Db 3202 AGCATCAATCAAAAGAGGGTTGGACATTGCATTTAGGATCATTTTCAACCCAA 3261
Qy 87 AspPheAlaSerAspAlaThrAsp-----ArgAsnGlyLysValTyrPheAla 104
Db 3262 AGTTTA-----GAACACGAGTAAAGAACTTTTAGACCAAAACTTATCTATAAGAC 3315
Qy 105 AspLysAsp---GlySerValThrLeuProThrGlnProValGlnGluPheLeuSer 123
Db 3316 CACAAGGATATGGCGGTATATTTAGCGAAGAT-----TATTGTAGC 3360
Qy 124 GlyHisValArgValArgProTyrLysGluLysProlleGlnAsnGln----- 139
Db 3361 GGCAACGTGAAGAAAGAACTCAAGAAAGTTAAAGAGGCATCAATCAAGCGGTGAGGA 3420
Qy 140 -----AlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAsp 157
Db 3421 TTAGAGGCTAATGCAAGATTTAGAGCTGATTATC-----CCTAAA 3462
Qy 158 AspAspPheArgProGlyLeuLysAspThrLysLeuLysLeuLysThrLeu----- 173
Db 3463 GAT-----TTGAAGCCACTGAATCATGGCTAATATCAACAGCCCTTGG 3507
Qy 173 ----- 173
Db 3508 ATACCCACTCAGTATTTAGAAGAGTTTTTAATGGAATTTAGCGCTAACCAATTATGAAAAG 3567
Qy 174 AlaIleGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsn 193
Db 3568 CAATACGGCGATAAATACAGATTTACCACTAAGCAATCTCAAGAGACATCAAAATA 3627
Qy 194 LysAsnHisProGly---TyrThrIleTyrGluArgAspSerSerIleValThrHisAsp 212
Db 3628 GAACACCTAAGCGGTGCTTATGAAGTTTTTGTAGAAACAATGAA----- 3672
Qy 213 AsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgValLysAsn 232
Db 3673 -----TTAAACGAGCTTTATGGTATCAGGCATATAAGAC 3705
Qy 233 ArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThr 252
Db 3706 AAGCCGCAATCTTATAAAGTGCCCTTTTGAAGCCCTTTTAAATAAAGTCTTAAACACAAG 3765
Qy 253 AspLeuIleSerGluLysTyrTyrValLeuLysLysGlyGluLysProTyrAspProPhe 272
Db 3766 GATTGT----- 3771
Qy 273 AspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAspThrAsnGluLeu 292
Db 3772 -----AGCGTTAAATACGCCCAAGTTGATCCTAATGACCC 3807
Qy 293 LeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyr 312
Db 3808 AAAAAAGAAATATTCATCCTGATGAAGAGCAAGCAATCTC----- 3849
Qy 313 AspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAsp 332
Db 3850 ---GCCACAAAGACAGAAATGAAGAAGCTTTTAAAGACTGGATTTATAAGAT 3906
Qy 333 TyrThrLeuThrGlyLysValGluAspAsnHisAspAspThr-----AsnArgIleIle 350
Db 3907 TATTCAAGAAGAACCCATTTAGAGCAAAATCTATAATGACACTTTCACCAACTCTGTTTA 3966
Qy 351 ThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyrHisLeuAla 368
```

```
Db 3967 AAAACCTATGATGCTCGCTCGCAATTAGAACTAGAGGGCTTTAACTACCATATCAGC 4020
RESULT 9
US-10-335-977-3326
; Sequence 3326, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 3326:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...7320
; SEQUENCE DESCRIPTION: SEQ ID NO: 3326:
US-10-335-977-3326
Alignment Scores:
Pred. No.: 0.00562 Length: 7320
Score: 138.50 Matches: 82
Percent Similarity: 36.93% Conservative: 65
Best Local Similarity: 20.60% Mismatches: 128
Query Match: 7.29% Indels: 123
DB: 16 Gaps: 16
US-09-940-235-2_COPY_16_383 (1-368) x US-10-335-977-3326 (1-7320)
Qy 21 LysPhePheGluIleAspLeuThr-----SerArg 30
Db 3214 AAAGACTTTGAAAAGAAATAATACCCCTAGAGCGCTAAATGCAAAACATAGAACCAAG 3273
Qy 31 ProAlaHisGlyGlyLysThr-----GluGlnGlyLeuSerProLysSerLys 46
Db 3274 CAAGCTCAAGCCAAAAAGAGCTCAAAATCTTTTGTGAAGGACTTTTAAACCTTAAAGAA 3333
Qy 47 PropheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLys 66
```

```
Db 3334 CTTATTACATAACGCT-----||||||-----AAAGAGCATTAATTGCA 3369
Qy 67 AlaileGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIle 86
Db 3370 AGCATCAATCAAAAGGGGTTGACCTGATTCATTAGGATCATTTTCACACCCAA 3429
Qy 87 AspPheAlaSerAspAlaThrIleThrAsp-----ArgAsnGlyLysValTyrPheAla 104
Db 3430 AGTTTA-----GAAACACGATTAAGAACTTTTAGAGCAAAACCTTATCTATAAGAC 3483
Qy 105 AspLysAsp---GlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSer 123
Db 3484 CACAGGATAATGGCGGTATATTATAGCGAACGAT-----TATTGAGC 3528
Qy 124 GlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGln-----139
Db 3529 GGCACGTAAGAAAGAACTCAAGAACTTAAGAAAGCCATCAATCAAGCGTGGAGGGA 3588
Qy 140 -----AlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAsp 157
Db 3589 TTAGAGGCTAATGTGAAGATTTAGAGCTGATTATC-----CCTAAA 3630
Qy 158 AspAspPheArgProGlyLeuLysAspThrLysLeuLysThrLeu-----173
Db 3631 GAT-----TTGAAAGCCATGAAATCATGCTAAATATCAACAGCCCTTGG 3675
Qy 173 -----173
Db 3676 ATACCACCTAGTATTAGAAAGTTTTAAATGGAATTTAGCGCTTACCATTAATAAG 3735
Qy 174 AlaileGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnSerIleLeuAsn 193
Db 3736 CAATACGGCGATAAATGACATATACCACTTAAGCAATCTCAAGAACACATCAAAATA 3795
Qy 194 LysAsnHisProGly---TyrThrIleTyrGluArgAspSerSerIleValThrHisAsp 212
Db 3796 GAACACCTAAGCGGTGCTTATGAAGTTTGTGTAGAAACAATGAA-----3840
Qy 213 AsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgValLysAsn 232
Db 3841 -----TTAAACGAGCTTTATGTTATGATCATGCGATAAAGAC 3873
Qy 233 ArgGluGlnAlaTyrArgIleAsnLysSerGlyLeuAsnGluGluIleAsnAsnThr 252
Db 3874 AAGCGCATTCCTATAAGTGCTTTTGAAGCCTTTTAAATAAAGTCTTAAACAACAAG 3933
Qy 253 AspLeuIleSerGluLysTyrTyrValLeuLysLysGlyLysProTyrAspProPhe 272
Db 3934 GATTG-----3939
Qy 273 AspArgSerHisLeuLysLeuPheThrIleLysTyrValAspValAspThrAsnGluLeu 292
Db 3940 -----AGCGTTAAATACGCCCAAGTTGATCCTAATGACCT 3975
Qy 293 LeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyr 312
Db 3976 AAAAAAGAAATATTCTCATCTGATGAAGAGCAAGCAATCTC-----4017
Qy 313 AspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAsp 332
Db 4018 ---GCCACACAAAGACAGAAATTTAAAGAGCTTTTAAAGACTGGATTTATAAGAT 4074
Qy 333 TyrThrLeuThrGlyLysValGluAspAsnHisAspAspThr-----AsnArgIleIle 350
Db 4075 TATTCAAGAAGAACCCATTTAGAGCAAACTATAATGACACTTTCAACAACCTTTGTTT 4134
Qy 351 ThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSerTyrHisLeuAla 368
Db 4135 AAAACCTATGATGGCTCGCAATTAGAACTAGAGGGCTTTAACTACCATATCAGC 4188
RESULT 10
US-10-398-221-2298
```

```
; Sequence 2298, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ IDS: 4025
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 2298
; LENGTH: 3714
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-2298

Alignment Scores:
Pred. No.: 0.218 Length: 3714
Score: 118.50 Matches: 86
Percent Similarity: 33.06% Conservative: 36
Best Local Similarity: 23.31% Mismatches: 104
Query Match: 6.23% Indels: 143
DB: 16 Gaps: 20

US-09-940-235-2_COPY_16_383 (1-368) x US-10-398-221-2298 (1-3714)
Qy 67 AlaileGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIle 86
Db 2569 AGCTTAGAGAACAGACTGTGACATATATCTATTCGAAAAATCCCGTT-----2616
Qy 87 AspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspLys 106
Db 2617 ---CCAGCTAAAGACATCAG-----GTACAATATACCGATGAG 2652
Qy 107 AspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisVal 126
Db 2653 GATGGATGGAATTAGCACCAACCCGAA-----ACATTATCAGGTAATGTT 2697
Qy 127 Arg-----127
Db 2698 GATGAAATATGTTTACAACAGCGAAGACTTTTACAGGATACGAATTAATTGAACACCG 2757
Qy 128 -----ValArgPro 130
Db 2758 AGCAATGCAGAAAGTTTAGCGAAATGCAACAACTGACCTATGTGTATCGCGC 2817
Qy 131 TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThrValGln 150
Db 2818 ATTAAGACGAGCCCAATT-----TTAGCTAAAGAAGTTACGGTGAACCTAC-----2862
Qy 151 PheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLysLeuLeu 170
Db 2863 -----CAGGACGAATTAGGAGCTTAAATTAAGTGAACCGAAGTATTG 2904
Qy 171 LysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnAlaGlnSer 190
Db 2905 ---ACTGGTGAATTTGGTGAACATACACACA-----2934
Qy 191 IleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAspSerSerIleValThr 210
Db 2935 ---GTGGCTAAACAATAGTAGTGGCTATACCTTAATTAAGAGCCCT-----ACT 2979
Qy 211 HisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgVal 230
Db 2980 AACCGCTAACGGGATTTTAAACGAG---AATCCTCAACCGGTTACCTATGTGTATCAATTA 3036
Qy 231 LysAsnArgGluGlnAlaTyrArgIleAsnLysSerGlyLeuAsnGluGluIleAsn 250
```



```
Qy      87 AspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaLeuLys 106
Db      1221543 ---CCAGTAAAGACATCAG-----GTACAATATACCGATGAG 1221578

Qy      107 AspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisVal 126
Db      1221579 GATGGATGGAATTAGCAACCAACCGAA-----ACATTATCAGGTAATGTT 1221623

Qy      127 Arg-----127

Db      1221624 GATGAAAATTATGTTACAACAGCGAAGACTTTTACAGGATACGAATTAATTGAACACCG 1221683

Qy      128 -----ValArgPro 130

Db      1221684 AGCAATGACAGAAAGAAAGTTTAGCGAAAATGCAAAACAGTGACCTATGTGTATCGCGG 1221743

Qy      131 TyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyrThrValGln 150
Db      1221744 ATTAAGCAGAGCCAATT-----TTAGCTAAAGAAGTTACGGTGAACCTAC----- 1221788

Qy      151 PheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThrLysLeuLeu 170
Db      1221789 -----CAGGACCAATTAGGAGCTTAAAAATTAGTCAAAACGGAAGTATTG 1221830

Qy      171 LysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeuLeuAlaGlnSer 190
Db      1221831 ---ACTGTGAAATTTGGTGAACATACACAACA-----1221860

Qy      191 IleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAspSerSerIleValThr 210
Db      1221861 ---GTGGCTAAACAATAGATGCTATACCTTAATTAAGAGCCT-----ACT 1221905

Qy      211 HisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgVal 230
Db      1221906 AACGCTAACGGATTTTAAACGAG---AATCCTCAAAACGGTTACCTATGTGTATCAATTA 1221962

Qy      231 LysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGluLeuLeu 250
Db      1221963 CAAATATGCCCAATAGCAGCAAAATATACCGTTAAACAC---TTGGATGAA-----AAC 1222013

Qy      251 AsnThrAspLeu-----IleSerGluLysTyrTyrVal 261
Db      1222014 AACACAGAGCTAGCACCAAGTGAAGTACTAAGCGGAATGTGATGAAGCATATACAACT 1222073

Qy      262 LeuLysLysGlyGluLysProTyrAsp-----ProPheAspArgSer----- 275
Db      1222074 AAACCAAAAGAAATAAAAGACTATTTCGCTAGTAAAAAGTACCAACTAATGCTAGCGGAAAA 1222133

Qy      276 -----HisLeuLys-----278

Db      1222134 TTCACGACAGAACGACAAACTGTGATTTATCTACTATAAAAGAACAGATATACCGACATCA 1222193

Qy      279 ---LeuPheThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGln 297
Db      1222194 TCTTATATTCTGTAATAATATGTAGATGAAACAGGAAAGACTGGCCATAGCAAGGTA 1222253

Qy      298 LeuLeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLys 317
Db      1222254 TTA-----AACGGAACATTAAATGATTCCTACGCAACTACAGCCAAG 1222295

Qy      318 AlaLysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeu----- 335
Db      1222296 GAA-----ATTAAGGCTATACGCTAGTGCAA 1222322

Qy      336 -----ThrGlyLysValGluAspAsnHisAspAspThrAsnArgIleIle 350
Db      1222323 AAGCCCGCTAATGCACTGGTAATTTACGGATCAGAA-----CAAACTATATAA 1222373

Qy      351 ThrValTyrMetGlyLysArgProGlu 359
Db      1222374 TATGTATATCGTCTAATAAAGCCGAA 1222400
```

```
RESULT 14
US-10-282-122A-8853
; Sequence 8853, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangau
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8853
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-8853

Alignment Scores:
Pred. No.: 0.0601 Length: 1323
Score: 118.00 Matches: 80
Percent Similarity: 38.27% Conservative: 57
Best Local Similarity: 22.35% Mismatches: 129
Query Match: 6.21% Indels: 92
DB: 16 Gaps: 17

US-09-940-235-2_COPY_16_383 (1-368) x US-10-282-122A-8853 (1-1323)

Qy      2 GlnLeuValValSerValAlaGlyThrValGlu-----12
Db      397 AAAATTGCAATTAGTGTAAATTAAACCTATTTCAGATTTTATCAAAATCTGTAGCCTTTA 456

Qy      13 -----GlyThrAsnGlnAspIleSerLeuLysPhe-----PheGluIleAspLeuThr 28
Db      457 GCATGGGGAATTAAATGCAATTGCCAATTTGCTATTCCGCTTGTGTTAAAGTAAATACCA 516

Qy      29 SerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLysSerLysProPhe 48
Db      517 -----CCGGAAGACAATATTACTTTTGTGATGATATTTTCAGCA 552

Qy      49 AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeuLysAlaIle 68
```

```

Db 553 GTGATGGAGGAGTCTCAAGCAGGTGTACTACAAAA----- 591
Qy 69 GlnGluGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluValIleAspPhe 88
Db 592 CAAGAACAT-----CACTTTATGAAACGTTATTTGAGCTTGAAGACGT 636
Qy 89 AlaSerAspAlaThrIleThrAspArgAsnGlyValPheAla-----AspLys 106
Db 637 ACAGTCCCATCAAGTATGACACACGCTGAAATGTTGTTTATTTACGCTTACAGCAT 696
Qy 107 AspGlySerVal-----ThrLeuProGlnProValGlnGluPheLeuLeuSer--- 123
Db 697 GAAGATAGCATTCGCCAAAACTGGCAGAAATATCTTATTCCAAATTTCTGGTTTGTAA 756
Qy 124 -----GlyHisValArgValArgProGlyLysGluLysProIle 136
Db 757 GAAATATTTGATCAACTGATGGCTATGTAGTCTAAGGATATCTTGTACGTATTATA 816
Qy 137 GlnAsnGlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnPro 156
Db 817 AACACCAATCT-----TTAACCCAGCTCAAT--- 843
Qy 157 AspAspAspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGly 176
Db 844 -----GAATCGATATTTCGTACGGTTCTTAATGATACCT 876
Qy 177 AspThrIleThrSerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsnHis 196
Db 877 GATACCTTAACTACTATCAAGATTAATCTCGATCGTCCGCTCAACAAAGAAAGTTGCT 936
Qy 197 ProGlyTyrThrIleTyrGluArgAspSerSerIleValHisAspAsnAspIlePhe 216
Db 937 GTTGTAATTAATAGTATCGCTGTGCTGGTGGGTAATTACGCTA---AGCGATATCATG 993
Qy 217 ArgThrIleLeu-----ProMetAspGlnGluPheThrTyrArgValLys 231
Db 994 ATTACTGTATGGGGATTTGGGTATACCCCTATGGAAGAGAA---CAGCAAAATTATATA 1050
Qy 232 AsnArgGluGlnAlaTyrArgIleAsn----- 240
Db 1051 CGTGATATAATCTTGGCTGATTTAGTGTAGTACACCTATTGAAGATCTTAAAGCAGCG 1110
Qy 241 ---LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyr 259
Db 1111 CTGGAATGGATGAATCACTACTGATGAAGAGATTTAAGAACATTAGCTGGTTTCATGATG 1170
Qy 260 TyrValLeuLysGlyGlyLysProTyrAspProPheAspArgSerHisLeuLysLeu 279
Db 1171 TACCGTTTACGCAAAATCCACGTCAGCGGACTTTGTTGAATTTGGTGGTTATAAA--- 1227
Qy 280 PheThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeuLeu 299
Db 1228 -----TTTGAAGTTGTTGATGATAGACCAT-----TTTAAATTTGATGCTATTATA 1272
Qy 300 ThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLys 317
Db 1273 GTAACA-----CGTGATTAGAAAAAATGATTTACCTTCTCCACCTGATGAA 1320

RESULT 15
US-10-425-114-3480
; Sequence 3480, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
```

```

; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3480
; LENGTH: 2332
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700266372_FLI
US-10-425-114-3480

Alignment Scores:
Pred. No.: 0.163 Length: 2332
Score: 117.00 Matches: 70
Percent Similarity: 37.82% Conservative: 48
Best Local Similarity: 22.44% Mismatches: 138
Query Match: 6.15% Indels: 56
DB: 16 Gaps: 12

US-09-940-235-2_COPY_16_383 (1-368) x US-10-425-114-3480 (1-2332)
Qy 69 GlnGluGlnLeuIleAlaAsnValHisSerAsnAspTyrPheGluValIleAspPhe 88
Db 1098 CAAGAACAGAAATGCTTAATGCGCAAGATCTTCTAAGGATATAATTTGAACATCTTCAAGC 1157
Qy 89 AlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAlaAspLysAspGly 108
Db 1158 CAGTATGTAGATTGACATCAGAAAAACAATGCTCTCAAAACTGAAATTTGAAGAACTGAAG 1217
Qy 109 SerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgVal 128
Db 1218 TCCAGACTCATAGATTGCGAGAGAACTCAAGAAATTTGTTATGGTTCAACATGTTGAAGAA 1277
Qy 129 ArgProTyrLysGluLysProIleGln----- 137
Db 1278 TGCCAAGTGGCTGAGATATAGATCAGAAGATTAGACTGCAAGCTGAAGTTCGCTCCTCC 1337
Qy 138 -----AsnGlnAlaLysSerValAspValIgluTyrThrValGlnPheThrProLeuAsn 155
Db 1338 AACATAAATCAGTTAGAAAAATTAGCTTCAGAACTACAGGAGAGAAATTCAGAGTTACTA 1397
Qy 156 ProAspAspAspPheArgProGlyLeuLysAspThrLysLeuLysThrLeuAlaIle 175
Db 1398 GAAGATTCTACCTTTGCTGAAAAATCAACAGCAAGAGTTGCTTCAAAAGATTTTGAAGCTA 1457
Qy 176 GlyAspThrIleThrSerGlnGluLeuAlaGlnAlaGlnSerIleLeu---AsnLys 194
Db 1458 GAATCA-----GATAATCAGGAACCTCTTAGGCAAAATGCAGTCCATATTATGGAAGAGAA 1511
Qy 195 AsnHisProGlyTyrThrIleTyr-----GluArgAspSerSerIleValThr 210
Db 1512 TCAAAATAGCTGAATCTCTGCATGGAGAGATAACTAAGCTGACAGCAGGTTGATACA 1571
Qy 211 HisAspAsnAspIle-----PheArgThrIleLeuProMetAspGlnGluPheThrTyr 228
Db 1572 CTTGAGAATCAGATCAACCAAGCTTCGTAGTGTCTG----- 1607
Qy 229 ArgValLysAsnArgGluGlnAlaTyr-----ArgIleAsnLysLysSerGlyLeuAsn 246
Db 1608 -----GATCAGAAGGAGCAACTCTACCTTTGTTCTGTACAAAAAGAGAGACTTTGGAG 1661
Qy 247 GluGlu---IleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGly 265
Db 1662 GAACAAAAATTCAGGTCGAGCAATTTGCTTTCTGCAACAGAGTGCACAACTTAGTGATCA 1721
Qy 266 GluLysProTyrAspProPheAspArg-----SerHisLeuLys 278
Db 1722 AAAAAACAGTATGATCTGATGCTTGAAGGCAAAAGATAGAGCTATCCAGCATTTGAAA 1781
Qy 279 LeuPheThrIleLysTyrValAspValAspThrAsnGluLeuLysSerGluGlnLeu 298
Db 1782 GAGCTATCTCTCAAAAAT---GACCAGGCGCATCAATGAAATCCGTAAAGAAATATGAACCT 1838
```

```
Qy 299 LeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAla 318
Db 1839 GAAAAGATTGAAATTACTAATGCTGAA-----AAAGAAAAGGCA 1877
Qy 319 LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLys 338
Db 1878 GAAAAGCTTCATAAGGGAATT-----GAAAACAAATGCAACGAAAAG 1919
Qy 339 ValGluAspAsnHisAspAspThrAsnArgIleIle 350
Db 1920 ATATCACAGACACAGCATGATTCTGAGAGGTATTG 1955
```

Search completed: November 6, 2004, 08:20:55
Job time : 1988.66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:42:10 ; Search time 37.8082 Seconds
(without alignments)
936.510 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901
Sequence: 1 SQLVSVAGTVEGTNQDISL.....ITVYMGKRPEGENASYHLA 368

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	100.0	440	1 A22801	streptokinase prec
2	1868.5	98.3	415	1 BZSO	streptokinase (EC
3	1859	97.8	440	2 S02723	streptokinase G pr
4	1726	90.8	440	2 S02724	streptokinase A pr
5	1597	84.0	440	2 S04168	streptokinase A pr
6	1590	83.6	414	2 JU0292	streptokinase - St
7	624.5	32.9	197	2 S53334	streptokinase - St
8	620	32.6	128	2 S77671	streptokinase A (E
9	569	29.9	128	2 S77680	streptokinase A (E
10	557	29.3	128	2 S77688	streptokinase A (E
11	552	29.0	128	2 S77679	streptokinase A (E
12	423	22.3	128	2 S77676	streptokinase A (E
13	419	22.0	128	2 S77673	streptokinase A (E
14	419	22.0	128	2 S77687	streptokinase A (E
15	417	21.9	128	2 S77682	streptokinase A (E
16	416	21.9	128	2 S77678	streptokinase A (E
17	408	21.5	128	2 S77677	streptokinase A (E
18	407	21.4	128	2 S77685	streptokinase A (E
19	401	21.1	128	2 S77683	streptokinase A (E
20	398	20.9	128	2 S77686	streptokinase A (E
21	397	20.9	128	2 S77684	streptokinase A (E
22	396	20.8	128	2 S77681	streptokinase A (E
23	390	20.5	128	2 S77674	streptokinase A (E
24	386	20.3	128	2 S77675	streptokinase A (E
25	140.5	7.4	2231	2 D71870	hypothetical prote
26	123	6.5	1019	2 T50251	hypothetical coile
27	120	6.3	1984	2 A44396	p-type cation tran
28	118.5	6.2	1237	2 AC1583	internalin protein
29	115.5	6.1	1512	2 T14883	hypothetical prote

30	115.5	6.1	1979	2	C71622	hypothetical prote
31	115.5	6.1	2256	2	AD1018	large repetitive p
32	115	6.0	1286	2	T16507	hypothetical prote
33	114.5	6.0	1435	2	S54697	DNA polymerase III
34	114.5	6.0	1435	2	C90596	hypothetical prote
35	114.5	6.0	3110	2	AC0116	probable virulence
36	113.5	6.0	785	2	D71653	cell surface anti
37	113.5	6.0	4688	2	F82885	hypothetical prote
38	112.5	5.9	558	2	F64402	vanadate-sensitiv
39	112	5.9	3394	2	T18501	hypothetical prote
40	110	5.8	967	2	S28428	phosphoenolpyruv
41	110	5.8	1745	2	A46431	tight junction-ass
42	110	5.8	13055	2	T16580	hypothetical prote
43	109.5	5.8	1220	1	DJBEC3	DNA-directed DNA p
44	109.5	5.8	1278	2	B70236	hypothetical prote
45	109	5.7	419	2	G72328	hypothetical prote

ALIGNMENTS

RESULT 1

A22801
streptokinase precursor - Streptococcus "equisimilis"
C;Species: Streptococcus "equisimilis"
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A22801
R;Malke, H.; Roe, B.; Ferretti, J.J.
Gene 34, 357-362, 1985

A;Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis H38
A;Reference number: A22801; MUID:95232082; PMID:2989113
A;Accession: A22801
A;Molecule type: DNA
A;Residues: 1-440 <MAL>
A;Cross-references: UNIPROT:P00779; GB:X72832; NID:g407876; PIDN:CAAS1351.1; PID:g407876
A;Experimental source: strain H46A
C;Genetics:
A;Gene: skc
C;Superfamily: streptokinase

Query Match 100.0%; Score 1901; DB 1; Length 440;
Best Local Similarity 100.0%; Pred. No. 4.9e-117;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	SQLVSVAGTVEGTNQDISLKFFFEIDLTSPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE	60
Db	42	SQLVSVAGTVEGTNQDISLKFFFEIDLTSPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE	101
Qy	61	KADLLKAIQELIANVHSNDYFEVIDFASDATITDRNGKVFADKDGSVTLPTQPVQEF	120
Db	102	KADLLKAIQELIANVHSNDYFEVIDFASDATITDRNGKVFADKDGSVTLPTQPVQEF	161
Qy	121	LLSGHVRVRYKPKPTQNAQSVDEVYVQFTPLNPDDDFRPGDKTLKLTLAGTGTIT	180
Db	162	LLSGHVRVRYKPKPTQNAQSVDEVYVQFTPLNPDDDFRPGDKTLKLTLAGTGTIT	221
Qy	181	SQLLAQAQSLNKNHPCYTIYERDSSIVTHDNDIFETILPMDQEFYRVKNREQAYRIN	240
Db	222	SQLLAQAQSLNKNHPCYTIYERDSSIVTHDNDIFETILPMDQEFYRVKNREQAYRIN	281
Qy	241	KKSGLNEEINTDLISEKYVVLKKGKPYOPDRSHLKLFTIKYVDVDTNELLKSEQULT	300
Db	282	KKSGLNEEINTDLISEKYVVLKKGKPYOPDRSHLKLFTIKYVDVDTNELLKSEQULT	341
Qy	301	ASERNLDFRDLYPDRAKALLYNLDAFGIMDYTLTGKVEDNHDHDTNR11TVYMGKRPEG	360
Db	342	ASERNLDFRDLYPDRAKALLYNLDAFGIMDYTLTGKVEDNHDHDTNR11TVYMGKRPEG	401
Qy	361	ENASYHLA 368	
Db	402	ENASYHLA 409	

RESULT 2

BZSO
streptokinase (EC 3.4.-.-) - Streptococcus sp.
C:Species: Streptococcus sp.
C:Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C:Accession: A00967
R:Jackson, K.W.; Tang, J.
Biochemistry 21, 6620-6625, 1982
A:Title: Complete amino acid sequence of streptokinase and its homology with serine protease
A:Reference number: A00967; MUID:83127125; PMID:6760891
A:Accession: A00967
A:Molecule type: protein
A:Residues: 1-415 <JAC>
A:Cross-references: UNIPROT:P00779
A:Note: 169-Asp and 181-Asp were also found
A:Note: this protein is not a protease, but it activates plasminogen by complexing with streptokinase
C:Superfamily: streptokinase
C:Keywords: hydrolase

Query Match 98.3%; Score 1868.5; DB 1; Length 415;
Best Local Similarity 98.9%; Pred. No. 6e-115; 2; Mismatches 1; Indels 1; Gaps 1;
Matches 365; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 16 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 75

QY 61 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
DB 76 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 135

QY 121 LLSGHVRVRPYKEKPIQONAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 180
DB 136 LLSGHVRVRPYKEKPIQONAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 195

QY 181 SQELLAQAQSILNKNHGGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
DB 196 SQELLAQAQSILNKNHGGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 255

QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 299
DB 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 315

QY 300 TASERNLDFRDLYPDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTRIIIVTMGRPE 359
DB 316 TASERNLDFRDLYPDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTRIIIVTMGRPE 375

QY 360 GENASYHLA 368
DB 376 GENASYHLA 384

RESULT 3

S02723
streptokinase G precursor - Streptococcus sp. (strain 19908)
C:Species: Streptococcus sp.
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S02723
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1262, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.
A:Reference number: S02723; MUID:89160265; PMID:2922269
A:Accession: S02723
A:Molecule type: DNA
A:Residues: 1-440 <WAL>
A:Cross-references: UNIPROT:P10519; EMBL:X13400; NID:g47095; PIDN:CAA31766.1; PID:g47096
C:Genetics:
A:Gene: skg
C:Superfamily: streptokinase
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-440/Product: streptokinase #status predicted <MAT>

Query Match 97.8%; Score 1859; DB 2; Length 440;

Best Local Similarity 97.8%; Pred. No. 2.7e-114;
Matches 360; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 42 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSKLFAFDGAMPKHLE 101

QY 61 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
DB 102 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 161

QY 121 LLSGHVRVRPYKEKPIQONAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 180
DB 162 LLSGHVRVRPYKEKPIQONAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 221

QY 181 SQELLAQAQSILNKNHGGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
DB 222 SQELLAQAQSILNKNHGGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 281

QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 300
DB 282 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 341

QY 301 ASERNLDFRDLYPDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTRIIIVTMGRPE 360
DB 342 ASERNLDFRDLYPDPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDNTRIIIVTMGRPE 401

QY 361 ENASYHLA 368
DB 402 ENASYHLA 409

RESULT 4

S02724
streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)
C:Species: Streptococcus pyogenes
C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 09-Jul-2004
C:Accession: S02724
R:Walter, F.; Siegel, M.; Malke, H.
Nucleic Acids Res. 17, 1261, 1989
A:Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes type
A:Reference number: S02724; MUID:89160264; PMID:2646590
A:Accession: S02724
A:Molecule type: DNA
A:Residues: 1-440 <WAL>
A:Cross-references: UNIPROT:P10520; EMBL:X13399; NID:g47435; PIDN:CAA31765.1; PID:g47436
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-440/Product: streptokinase #status predicted <MAT>

Query Match 90.8%; Score 1726; DB 2; Length 440;

Best Local Similarity 89.1%; Pred. No. 1.4e-105;
Matches 328; Conservative 20; Mismatches 20; Indels 0; Gaps 0;

QY 1 SOLVSVAGTVEGTNQDISLKFEIDLTSRPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
DB 42 SOLVSMAGIVEGTDKKVFINFEIDLTSQPAHGGKTEQGLSPKSPKPFATDNGAMPKHLE 101

QY 61 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 120
DB 102 KADLLKAIQELIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQPVQEF 161

QY 121 LLSGHVRVRPYKEKPIQONAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 180
DB 162 LVKGHVRVRPYKEKPIQONAKSVDEYTVQFTPLNPDFFRPGDKDTLLKTLAIGDTIT 221

QY 181 SQELLAQAQSILNKNHGGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 240
DB 222 SQELLAQAQSILNKNHGGTYIYERDSSIVTHDNDIFRTILPMDQEFYRVKNREQAYRIN 281

QY 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 300

Db 282 PKTGIEKTNNDYLVSEKYYVLLKQSGKPYDPFDRSHLKLFTIKYVDVNTNELLKSQLLT 341
Qy 301 ASERNLDFRDLDPDRKAKLLYNLDLAFGIMDYTLTGKVEDNDDTNRITVTVMGRPEG 360
Db 342 ASEGNDLFRDLDPDRKAKLLYNLDLAFIMTYTLTGKVEDNDDKNNRVVTVMGRPKG 401
Qy 361 ENASYHLA 368
Db 402 AKGSYHLA 409
RESULT 5
S04168
streptokinase A precursor - Streptococcus pyogenes (strain NZ131)
C:Species: Streptococcus pyogenes
C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: S04168
R:Huang, T.T.; Malke, H.; Ferretti, J.J.
Mol. Microbiol. 3, 197-205, 1989
A:Title: The streptokinase gene of group A streptococci: cloning, expression in Escherichia coli and characterization of the gene product.
A:Reference number: S04168; MUID:89343623; PMID:2668686
A:Accession: S04168
A:Molecule type: DNA
A:Residues: 1-440 <HUA>
C:Genetics:
A:Gene: ska
C:Superfamily: streptokinase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-440/Product: streptokinase A #status predicted <MAT>
Query Match 84.0%; Score 1597; DB 2; Length 440;
Best Local Similarity 84.2%; Pred. No. 3.8e-97; Indels 0; Gaps 0;
Matches 310; Conservative 22; Mismatches 36; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 42 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSPFATDKGAMSHKLE 101
Qy 61 KADLLKAIQEQLIANVHNSNDYFEVIDFASDATITDRNGKVPADKDGSVTLTPQVQEF 120
Db 102 KADLLKAIQEQLIANVHNSNDYFEVIDFASDATITDRNGKVPADKDDSVTLTPQVQEF 161
Qy 121 LLSGHVRVPYKEKPTQNOAKSVVDVYTFQPLNPDDEFPGDKTLKTLAIGDTIT 180
Db 162 LLSGHVRVPYKEKPTQNOAKSVVDVYTFQPLNPDDEFPLRNQYHLTTLAVGDSLS 221
Qy 181 SOELLAQAQSIILKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 240
Db 222 SOELLAQAQSIILKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 281
Qy 241 KKSGLNEEINNTDLISEKYYVLLKQSGKPYDPFDRSHLKLFTIKYVDVNTNELLKSQLLT 300
Db 282 SKTGIEEKTNTDLISEKYYVLLKQSGKPYDPFDRSHLKLFTIKYVDVNTNELLKSQLLT 341
Qy 301 ASERNLDFRDLDPDRKAKLLYNLDLAFGIMDYTLTGKVEDNDDTNRITVTVMGRPEG 360
Db 342 ASERNLDFRDLDPDRKAKLLYNLDLAFGIMDYTLTGKVEDNDDTNRITVTVMGRPEG 401
Qy 361 ENASYHLA 368
Db 402 ENASYHLA 409
RESULT 6
JU0292
streptokinase - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A43867; JU0292
R:Ohkuni, H.; Todome, Y.; Suzuki, H.; Mizuse, M.; Kotani, N.; Horiuchi, K.; Shikama, N.;
Infect. Immun. 60, 278-283, 1992

A:Title: Immunochemical studies and complete amino acid sequence of the streptokinase f
A:Reference number: A43867; MUID:92104686; PMID:1370275
A:Accession: A43867
A:Molecule type: protein
A:Residues: 1-414 <OHK>
A:Cross-references: UNIPROT:Q57391
A:Experimental source: M type 12 strain A374
A:Note: sequence extracted from NCBI backbone (NCBIP:74592)
C:Superfamily: streptokinase
Query Match 83.6%; Score 1590; DB 2; Length 414;
Best Local Similarity 84.0%; Pred. No. 1e-96; Indels 0; Gaps 0;
Matches 309; Conservative 23; Mismatches 36; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVEGTNQDISLKFFEIDLTSRPAHGGKTEQGLSPKSPFATNSAMPHKLE 75
Qy 61 KADLLKAIQEQLIANVHNSNDYFEVIDFASDATITDRNGKVPADKDGSVTLTPQVQEF 120
Db 76 KADLLKAIQEQLIANVHNSNDYFEVIDFASDATITDRNGKVPADRDSVTLPQVQEF 135
Qy 121 LLSGHVRVPYKEKPTQNOAKSVVDVYTFQPLNPDDEFPGDKTLKTLAIGDTIT 180
Db 136 LLSGHVRVPYKEKPTQNOAKSVVDVYTFQPLNPDDEFPLSKERYHLTTLAVGDSLS 195
Qy 181 SOELLAQAQSIILKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 240
Db 196 SOELLAQAQSIILKNHGPYTIYERDSSIVTHDNDIFRTILPMDQEFYRKNREQAYRIN 255
Qy 241 KKSGLNEEINNTDLISEKYYVLLKQSGKPYDPFDRSHLKLFTIKYVDVNTNELLKSQLLT 300
Db 256 SKTGIVEKTNTDLISEKYYVLLKQSGKPYDPFDRSHLKLFTIKYVDVNTNELLKSQLLT 315
Qy 301 ASERNLDFRDLDPDRKAKLLYNLDLAFGIMDYTLTGKVEDNDDTNRITVTVMGRPEG 360
Db 316 ASERNLDFRDLDPDRKAKLLYNLDLAFGIMDYTLTGKVEDNDDTNRITVTVMGRPEG 375
Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383
RESULT 7
S53334
streptokinase - Streptococcus sp.
C:Species: Streptococcus sp.
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S53334
R:Shi, G.Y.; Chang, B.I.; Chen, S.M.; Wu, D.H.; Wu, H.L.
Biochem. J. 304, 235-241, 1994
A:Title: Function of streptokinase fragments in plasminogen activation.
A:Reference number: S53334; MUID:95091634; PMID:7998939
A:Accession: S53334
A:Molecule type: protein
A:Status: preliminary
A:Residues: 1-197 <SHI>
A:Cross-references: UNIPROT:Q7M115
Query Match 32.9%; Score 624.5; DB 2; Length 197;
Best Local Similarity 47.8%; Pred. No. 5.5e-34;
Matches 154; Conservative 2; Mismatches 3; Indels 163; Gaps 8;
Qy 42 SPKSPFATDSGAMSHKLEKADLLKAIQEQLIANVHNSNDYFEVIDFASDATITDRNGKV 101
Db 12 SVNSKPFATDSGAMSHKLEKADLLKAIQEQIL-----NGKV 47
Qy 102 YFADKDGSVTLTPQVQEFLLSGHVRVPYKEKPTQNOAKSVVDVYTFQPLNPDDEF 161
Db 48 YFADKDGSVT-----EKPTQNOAKSVVDV----- 71
Qy 162 PGLKDTKLLKTLAIGDTITTSQELLAQAQSIILKNHGPYTIYERDSSIVTHDNDIFRTILP 221

```
Db 72 -----NHGTYTYERDSSIV-----TILP 90
Qy 222 MDQETVRVKNREQAYRINKKSGLNNEINNTDLISEKYVVLKKGKPYDPDRSHLKLFT 281
Db 91 MDQETTY-----SGLNEEINNTDLIS----- 111
Qy 282 IKYVDVDTNELKSEQLLTASERNLDFRDLDPDRKAKLLYNLDAFGIMDYTLTGKVED 341
Db 112 --YVDVDTNELKSEQLLTASERN-----DLYDPRDKAKLLYNLDAFGIM----- 155
Qy 342 NHDNTNRITVYMGKRPGEENA 363
Db 156 -----IITVYMGKRPGEENA 170

RESULT 8
S77671
streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
C:Accession: S77671; S77672
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77671
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: EMBL:U25853; NID:g818908; PIDN:AAA85729.1; PID:g818909
A:Experimental source: strain ET1/M1
A:Note: allele 2
A:Accession: S77672
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KA2>
A:Cross-references: EMBL:U25854; NID:g818910; PIDN:AAA85730.1; PID:g818911
A:Experimental source: strain E2/M3
A:Note: allele 3
C:Genetics:
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Query Match 32.6%; Score 620; DB 2; Length 128;
Best Local Similarity 92.2%; Pred. No. 5.9e-34;
Matches 118; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 126 VVRVYKKEPIQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 185
Db 1 VVRVYKKEPVQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 60
Qy 186 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 245
Db 61 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 120
Qy 246 NEEINNTD 253
Db 121 KEKTNNTD 128

RESULT 9
S77680
streptokinase A (EC 3.4.-.-) (allele 11) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77680
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77680
A>Status: translation not shown
A:Keywords: hydrolase; plasminogen activator; virulence

Query Match 32.6%; Score 620; DB 2; Length 128;
Best Local Similarity 92.2%; Pred. No. 5.9e-34;
Matches 118; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 126 VVRVYKKEPIQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 185
Db 1 VVRVYKKEPVQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 60
Qy 186 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 245
Db 61 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 120
Qy 246 NEEINNTD 253
Db 121 KEKTNNTD 128

RESULT 10
S77688
streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77688
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77688
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54695; EMBL:U25870; NID:g818942; PIDN:AAA85741.1; PID:g818943
A:Experimental source: strain ET76/M72
C:Genetics:
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Query Match 29.3%; Score 557; DB 2; Length 128;
Best Local Similarity 83.6%; Pred. No. 7.8e-30;
Matches 107; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 126 VVRVYKKEPIQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 185
Db 1 VVRVYKKEPIQPAKSVDIRYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 60
Qy 186 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 245
Db 61 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 120
Qy 246 NEEINNTD 253
Db 121 KEKTNNTD 128

RESULT 11
S77679
streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77679
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
```

```
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54687; EMBL:U25862; NID:g818926; PIDN:AAA85738.1; PID:g818929
A:Experimental source: strain ET51/M17
C:Genetics:
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Query Match 29.9%; Score 569; DB 2; Length 128;
Best Local Similarity 85.9%; Pred. No. 1.3e-30;
Matches 110; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Qy 126 VVRVYKKEPIQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 185
Db 1 VVRVYKKEPIQPAKSVDIRYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 60
Qy 186 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 245
Db 61 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGQ 120
Qy 246 NEEINNTD 253
Db 121 EEKTNNTD 128

RESULT 10
S77688
streptokinase A (EC 3.4.-.-) (allele 19) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77688
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77688
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <KAP>
A:Cross-references: UNIPROT:Q54695; EMBL:U25870; NID:g818942; PIDN:AAA85741.1; PID:g818943
A:Experimental source: strain ET76/M72
C:Genetics:
C:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Query Match 29.3%; Score 557; DB 2; Length 128;
Best Local Similarity 83.6%; Pred. No. 7.8e-30;
Matches 107; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 126 VVRVYKKEPIQNOAKSVDEVYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 185
Db 1 VVRVYKKEPIQPAKSVDIRYTVQFTPLNPDDEFPGKDKTKLLTKLAIGDTITSQELL 60
Qy 186 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 245
Db 61 AQAQSIILNKHGPGTYTYERDSSIVTHDNDIFRILPMDQEFYVRVKNREQAYRINKKSGL 120
Qy 246 NEEINNTD 253
Db 121 KEKTNNTD 128

RESULT 11
S77679
streptokinase A (EC 3.4.-.-) (allele 10) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: S77679
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
```

A;Reference number: S77671; MUID:96037795; PMID:7565111
A;Accession: S77679
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-128 <KAP>
A;Cross-references: UNIPROT:Q54686; EMBL:U25861; NID:g818924; PIDN:AAA85737.1; PID:g8189
A;Experimental source: strain ET50/M43
C;Genetics:
A;Gene: ska
C;Superfamily: streptokinase
C;Keywords: hydrolase; plasminogen activator; virulence

Query Match 29.0%; Score 552; DB 2; Length 128;
Best Local Similarity 82.8%; Pred. No. 1.7e-29;
Matches 106; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 126 VVRVPYKEPIQNAQKSVDEYTVQFTPLNPDGKDTKLLKTLAGDITTSOELL 185
DB 1 VVRVPYKEPIQNAQKSVDIRAVQFTPLNPDGKDTKLLKTLAGDITTSOELL 60
QY 186 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
DB 61 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
QY 246 NEEINNTD 253
DB 121 KEKINNTD 128

RESULT 12

S77676
streptokinase A (EC 3.4.-.-) (allele 7) - Streptococcus pyogenes (fragment)
C;Species: Streptococcus pyogenes
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
A;Accession: S77676
R;Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A;Reference number: S77671; MUID:96037795; PMID:7565111
A;Accession: S77676
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-128 <KAP>
A;Cross-references: UNIPROT:Q54683; EMBL:U25858; NID:g818918; PIDN:AAA85734.1; PID:g8189
A;Experimental source: ET16/M66
C;Genetics:
A;Gene: ska
C;Superfamily: streptokinase
C;Keywords: hydrolase; plasminogen activator; virulence

Query Match 22.3%; Score 423; DB 2; Length 128;
Best Local Similarity 66.4%; Pred. No. 4.6e-21;
Matches 85; Conservative 12; Mismatches 31; Indels 0; Gaps 0;

QY 126 VVRVPYKEPIQNAQKSVDEYTVQFTPLNPDGKDTKLLKTLAGDITTSOELL 185
DB 1 VVRVPYQPKAVHNSAERVNVNVEVSFVSETGDLDFPSLRDHYHLTTTLAGVDSLSSQELA 60
QY 186 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
DB 61 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
QY 246 NEEINNTD 253
DB 121 KEKINNTD 128

RESULT 13

S77673
streptokinase A (EC 3.4.-.-) (allele 4) - Streptococcus pyogenes (fragment)
C;Species: Streptococcus pyogenes
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
A;Accession: S77673

R;Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A;Reference number: S77671; MUID:96037795; PMID:7565111
A;Accession: S77673
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-128 <KAP>
A;Cross-references: UNIPROT:Q54680; EMBL:U25855; NID:g818912; PIDN:AAA85731.1; PID:g8189
A;Experimental source: strainn ET4/M2
C;Genetics:
A;Gene: ska
C;Superfamily: streptokinase
C;Keywords: hydrolase; plasminogen activator; virulence

Query Match 22.0%; Score 419; DB 2; Length 128;
Best Local Similarity 65.6%; Pred. No. 8.4e-21;
Matches 84; Conservative 13; Mismatches 31; Indels 0; Gaps 0;

QY 126 VVRVPYKEPIQNAQKSVDEYTVQFTPLNPDGKDTKLLKTLAGDITTSOELL 185
DB 1 VVRVPYQPKAVHNSAERVNVNVEVSFVSETGDLDFPSLRDHYHLTTTLAGVDSLSSQELA 60
QY 186 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
DB 61 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
QY 246 NEEINNTD 253
DB 121 KEKINNTD 128

RESULT 14

S77687
streptokinase A (EC 3.4.-.-) (allele 18) - Streptococcus pyogenes (fragment)
C;Species: Streptococcus pyogenes
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
A;Accession: S77687
R;Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser,
Mol. Microbiol. 16, 509-519, 1995
A;Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus
A;Reference number: S77671; MUID:96037795; PMID:7565111
A;Accession: S77687
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-128 <KAP>
A;Cross-references: UNIPROT:Q54694; EMBL:U25869; NID:g818940; PIDN:AAA85740.1; PID:g8189
A;Experimental source: strain ET74/M46
C;Genetics:
A;Gene: ska
C;Superfamily: streptokinase
C;Keywords: hydrolase; plasminogen activator; virulence

Query Match 22.0%; Score 419; DB 2; Length 128;
Best Local Similarity 64.8%; Pred. No. 8.4e-21;
Matches 83; Conservative 15; Mismatches 30; Indels 0; Gaps 0;

QY 126 VVRVPYKEPIQNAQKSVDEYTVQFTPLNPDGKDTKLLKTLAGDITTSOELL 185
DB 1 VVRVPYQPKAVHNSAERVNVNVEVSFVSETGDLDFPSLRDHYHLTTTLAGVDSLSSQELA 60
QY 186 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
DB 61 AQAQSLNKNHPCGTYIYERDSSIVTHDNDIFRILPMDQEFYVKNRQAYRINKKSGL 245
QY 246 NEEINNTD 253
DB 121 KEKINNTD 128

RESULT 15

S77682
streptokinase A (EC 3.4.-.-) (allele 13) - Streptococcus pyogenes (fragment)

C:Species: Streptococcus pyogenes
 C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
 C:Accession: S77682
 R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser, M.O. Microbiol. 16, 509-519, 1995
 A:Title: Molecular population genetic analysis of the streptokinase gene of Streptococcus pyogenes
 A:Reference number: S77671; MUID:96037795; PMID:7565111
 A:Accession: S77682
 A>Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-128 <KAP>
 A:Cross-references: UNIPROT:Q54689; EMBL:U25864
 A:Experimental source: strain ET55/M25
 C:Genetics:
 A:Gene: ska
 C:Superfamily: streptokinase
 C:Keywords: hydrolase; plasminogen activator; virulence

Query Match 21.9%; Score 417; DB 2; Length 128;
 Best Local Similarity 65.6%; Pred. No. 1.1e-20;
 Matches 84; Conservative 13; Mismatches 31; Indels 0; Gaps 0;
 QY 126 VVRVPYKEKPIQNAKSVDEYTVQFTPLNPDDEFGLKDTKLXLAIGDTITTSQELL 185
 Db 1 VVRVPYQPKAVHNSAERVNVNVEVSFVSETGNLDFTPSLKERYHLTTLAVGDSLSSQELA 60
 QY 186 AQAQSIILNKNHPGYTIYERDSSIVTHDNDIFRTILPMDQEFYVRVKNRQAYRINKKSGL 245
 Db 61 AIAQFILSKHPDYIITKRDSIVTHDNDIFRTILPMDQEFYVRVKNRQAYRINKKSGQ 120
 QY 246 NEEINNTD 253
 Db 121 KEKINNTD 128

Search completed: November 3, 2004, 23:55:41
 Job time : 39.8082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:54:46 ; Search time 134.849 Seconds
(without alignments)
884.776 Million cell updates/sec

Title: US-09-940-235-2_COPY_16_383

Perfect score: 1901

Sequence: 1 SOLVSVAGTVEGTNQDISL.....IITVYMGKPEGENASYHLA 368

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1901	100.0	413	15	US-10-360-101-264
2	1901	100.0	414	10	US-09-940-235-2
3	1875	98.6	414	14	US-10-300-215-252
4	1864	98.1	414	14	US-10-300-215-253
5	1845	97.1	413	9	US-09-919-703-12
6	140.5	7.4	2384	15	US-10-335-977-8087
7	138.5	7.3	2440	15	US-10-335-977-8088
8	118	6.2	441	15	US-10-328-122A-45037
9	117	6.2	777	15	US-10-425-114-59217
10	117	6.2	798	15	US-10-425-114-66924
11	115.5	6.1	5559	15	US-10-282-122A-75177
12	115	6.0	21	9	US-09-919-703-7
13	113.5	6.0	561	15	US-10-282-122A-71575
					Sequence 264, App
					Sequence 2, Appli
					Sequence 252, App
					Sequence 253, App
					Sequence 12, Appl
					Sequence 8087, Ap
					Sequence 8088, Ap
					Sequence 45037, A
					Sequence 59217, A
					Sequence 66924, A
					Sequence 75177, A
					Sequence 7, Appli
					Sequence 71575, A

ALIGNMENTS

RESULT 1

US-10-360-101-264

; Sequence 264, Application US/10360101

; Publication No. US20040009550A1

; GENERAL INFORMATION:

; APPLICANT: Moll, Gert N.

; APPLICANT: Leenhouts, Cornelis J.

; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way

; FILE REFERENCE: 2183-5673

; CURRENT APPLICATION NUMBER: US/10/360,101

; PRIOR FILING DATE: 2003-02-07

; PRIOR APPLICATION NUMBER: EP 02077060.8

; NUMBER OF SEQ ID NOS: 309

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 264

; LENGTH: 413

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: sequence of streptokinase

; US-10-360-101-264

Query Match 100.0%; Score 1901; DB 15; Length 413;

Best Local Similarity 100.0%; Pred. No. 2.5e-144;

Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKFFIEDLTSRPAHGKTEQGLSPKSPFATDSGAMSHKLE 60

Db 16 SOLVSVAGTVEGTNQDISLKFFIEDLTSRPAHGKTEQGLSPKSPFATDSGAMSHKLE 75

Qy 61 KADLLKAIQBLIANVHNSDDYFVDFASDAITDRNGKVYFADKDGSTLPTQVQEF 120

Db 76 KADLLKAIQBLIANVHNSDDYFVDFASDAITDRNGKVYFADKDGSTLPTQVQEF 135

Qy 121 LLSGHVVRVRYKEKPIQNAQKSDVVEYVTFQTLNPDPPDFRPGIKDTKLKLTIAIGTIT 180

Sequence 76865, A

Sequence 129902, A

Sequence 126450, A

Sequence 46664, A

Sequence 20, Appl

Sequence 29, Appl

Sequence 2471, Ap

Sequence 240347, A

Sequence 53254, A

Sequence 53353, A

Sequence 45, Appl

Sequence 45, Appl

Sequence 71900, A

Sequence 52796, A

Sequence 72314, A

Sequence 51864, A

Sequence 53190, A

Sequence 82, Appl

Sequence 106, App

Sequence 123, App

Sequence 4, Appli

Sequence 4, Appli

Sequence 3, Appli

Sequence 3, Appli

Sequence 3, Appli

Sequence 60608, A

Sequence 72134, A

Sequence 181381, A

Db 136 LLSGHRVRPYKEKPIQNAKSDVEYTVQFTPLNPDDEFGLKDTLLKTLAIGDTIT 195
Qy 181 SOELLAQAQSILNKHHPGVTIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
Db 196 SOELLAQAQSILNKHHPGVTIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 255
Qy 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Qy 301 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 360
Db 316 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 375
Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 2

US-09-940-235-2
; Sequence 2, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-09-940-235-2

Query Match 100.0%; Score 1901; DB 10; Length 414;
Best Local Similarity 100.0%; Pred. No. 2.5e-144;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPFATDSCAMSHKLE 60
Db 16 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPFATDSCAMSHKLE 75
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 120
Db 76 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 135
Qy 121 LLSGHRVRPYKEKPIQNAKSDVEYTVQFTPLNPDDEFGLKDTLLKTLAIGDTIT 180
Db 136 LLSGHRVRPYKEKPIQNAKSDVEYTVQFTPLNPDDEFGLKDTLLKTLAIGDTIT 195
Qy 181 SOELLAQAQSILNKHHPGVTIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
Db 196 SOELLAQAQSILNKHHPGVTIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 255
Qy 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300

Db 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Qy 301 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 360
Db 316 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 375
Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 3

US-10-300-215-252
; Sequence 252, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 252
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Streptococcus equisimilis
US-10-300-215-252

Query Match 98.6%; Score 1875; DB 14; Length 414;
Best Local Similarity 98.6%; Pred. No. 3.1e-142;
Matches 363; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPFATDSCAMSHKLE 60
Db 16 SOLVSVAGTVEGTNQDISLKPFEDLTSRPAHGGKTEQGLSPKSPFATDSCAMSHKLE 75
Qy 61 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 120
Db 76 KADLLKAIQEOQLIANVHNSNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTQVQEF 135
Qy 121 LLSGHRVRPYKEKPIQNAKSDVEYTVQFTPLNPDDEFGLKDTLLKTLAIGDTIT 180
Db 136 LLSGHRVRPYKEKPIQNAKSDVEYTVQFTPLNPDDEFGLKDTLLKTLAIGDTIT 195
Qy 181 SOELLAQAQSILNKHHPGVTIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 240
Db 196 SOELLAQAQSILNKHHPGVTIYERDSSIIVTHNDIFRTILPMDQEFYRVKREQAYRIN 255
Qy 241 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 256 KKSGLNEEINNTDLISEKYVVLKKGKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Qy 301 ASERNLDFRDLPRDKAKLLYNNLDAGFIMDYTLTGKVEDNHDHDTNRIITVYMGKREPEG 360

Db 316 ASERNLDFRDLYDPDRKAKLLYNNLDLAFGIMDYTLTGKVEDNHDNTRIIITVYMGKRPEG 375
Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 4
US-10-300-215-253
; Sequence 253, Application US/10300215
; Publication No. US20030153043A1
; GENERAL INFORMATION:
; APPLICANT: CARR, Francis Joseph
; APPLICANT: ADAIR, Fiona Suzanne
; APPLICANT: HAMILTON, Anita Anne
; APPLICANT: CARTER, Graham
; TITLE OF INVENTION: METHOD FOR THE PRODUCTION OF
; TITLE OF INVENTION: NON-IMMUNOGENIC PROTEINS
; FILE REFERENCE: MER-104-Con.1
; CURRENT APPLICATION NUMBER: US/10/300,215
; CURRENT FILING DATE: 2002-11-20
; PRIOR APPLICATION NUMBER: US 09/438,136
; PRIOR FILING DATE: 1999-11-10
; PRIOR APPLICATION NUMBER: WO PCT/GB98/01473
; PRIOR FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: GB 9710480.6
; PRIOR FILING DATE: 1997-05-21
; PRIOR APPLICATION NUMBER: GB 9716197.0
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: GB 9725270.4
; PRIOR FILING DATE: 1997-11-28
; PRIOR APPLICATION NUMBER: GB 9807751.4
; PRIOR FILING DATE: 1998-04-14
; PRIOR APPLICATION NUMBER: US 60/067,235
; PRIOR FILING DATE: 1997-12-02
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Modified strep protein
US-10-300-215-253

Query Match 98.1%; Score 1864; DB 14; Length 414;
Best Local Similarity 98.1%; Pred. No. 2.3e-141;
Matches 361; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 1 SOLVSVAGTVGEGTQDLSLKPFEDLTSPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTVGEGTQDLSLKPFEDLTSPAHGGKTEQGLSPKSPKPFATDSGAMPHKLE 75
Qy 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 120
Db 76 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 135
Qy 121 LLSGHVRVRYPEKPIQNAQSVDEYTVQFTPLNPDHDIPTILPMDQEFYRKNREQAYRN 180
Db 136 LLSGHVRVRYPEKPIQNAQSVDEYTVQFTPLNPDHDIPTILPMDQEFYRKNREQAYRN 195
Qy 181 SQELLAAQASILNKNHPGYTIYERDSSIVTHNDIPTILPMDQEFYRKNREQAYRN 240
Db 196 SQELLAAQASILNKNHPGYTIYERDSSIVTHNDIPTILPMDQEFYRKNREQAYRN 255
Qy 241 KKSGLNEEINNTDLISEKYVLLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 300
Db 256 KKSGLNEEINNTDLISEKYVLLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLLT 315
Qy 301 ASERNLDFRDLYDPDRKAKLLYNNLDLAFGIMDYTLTGKVEDNHDNTRIIITVYMGKRPEG 360
Db 316 ASERNLDFRDLYDPDRKAKLLYNNLDLAFGIMDYTLTGKVEDNHDNTRIIITVYMGKRPEG 375

Qy 361 ENASYHLA 368
Db 376 ENASYHLA 383

RESULT 5
US-09-919-703-12
; Sequence 12, Application US/09919703
; Patent No. US20020165129A1
; GENERAL INFORMATION:
; APPLICANT: Kystal, Gerald
; APPLICANT: Rabkin, Simon W.
; TITLE OF INVENTION: Peptides and Their Use to Ameliorate
; TITLE OF INVENTION: Cell Death
; FILE REFERENCE: 50216/003004
; CURRENT APPLICATION NUMBER: US/09/919,703
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 09/294,457
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: US 08/759,599
; PRIOR FILING DATE: 1996-12-05
; PRIOR APPLICATION NUMBER: US 60/008,233
; PRIOR FILING DATE: 1995-12-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-12

Query Match 97.1%; Score 1845; DB 9; Length 413;
Best Local Similarity 98.6%; Pred. No. 7.8e-140;
Matches 362; Conservative 1; Mismatches 2; Indels 2; Gaps 2;
Qy 1 SOLVSVAGTVGEGTQDLSLKPFEDLTSPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 60
Db 16 SOLVSVAGTV-GTNQDLSLKPFEDLTSPAHGGKTEQGLSPKSPKPFATDSGAMSHKLE 74
Qy 61 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 120
Db 75 KADLLKAIQEQLIANVHNSDDYFEVIDFASDATITDRNGKVPYFADKDGSVTLTPQVQEF 134
Qy 121 LLSGHVRVRYPEKPIQNAQSVDEYTVQFTPLNPDHDIPTILPMDQEFYRKNREQAYRN 180
Db 135 LLSGHVRVRYPEKPIQNAQSVDEYTVQFTPLNPDHDIPTILPMDQEFYRKNREQAYRN 194
Qy 181 SQELLAAQASILNKNHPGYTIYERDSSIVTHNDIPTILPMDQEFYRKNREQAYRN 240
Db 195 SQELLAAQASILNKNHPGYTIYERDSSIVTHNDIPTILPMDQEFYRKNREQAYRN 254
Qy 241 KKSGLNEEINNTDLISEKYVLLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 299
Db 255 KKSGLNEEINNTDLISEKYVLLKGEKPYDPDRSHLKLFTIKYVDVDTNELLKSEQLL 314
Qy 300 TASERNLDFRDLYDPDRKAKLLYNNLDLAFGIMDYTLTGKVEDNHDNTRIIITVYMGKRPE 359
Db 315 TASERNLDFRDLYDPDRKAKLLYNNLDLAFGIMDYTLTGKVEDNHDNTRIIITVYMGKRPE 374
Qy 360 GENASYH 366
Db 375 GENASYH 381

RESULT 6
US-10-335-977-8087
; Sequence 8087, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

185	Db	VMDAGAQAQVULQK-----QEH-----HFTENVPELEERTVPSMTHTRNVVVYFTLNH	232
107	Qy	DGSV--TLTPQPVQEFLLS-----GHVRVPYKEKPIQIOAKSVDEVYTVQFTPLNP	156
233	Db	EDSIROKLAEPYPSKFLVCNENIDQVIGYVDKDLVRLNQS-----LTQLN--	281
157	Qy	DDDFRPGDKTKLLKTLAIGDTITSQELLAAQOSILNKHNGYTYIERDSSIVTHDNDIF	216
282	Db	-----ESTITVLMIPTTLSELLDRFRSTKERFAVVINEYALVWGVTIL--SDIM	331
217	Qy	RTIL-----PMDQEFYTVKVKRQOAYRN-----KKSGLNEEINNTDLISEKY	259
332	Db	ITVMGDWVTPMEEB--QQIIKRDNNNSWLDIGSTPTEDLKHALEMDELPEENYETVLAGPMM	390
260	Qy	YVLKKGBKPYDPFDRSHKLFTIKYVDVDTNELLKSEQLLTASERNLDFRDLYPDPRK	317
391	Db	YLRKIPRPADVFBEFGVK---FEVVDVH---FKIDOLLVT--RVLEKNLPPSPDE	440

2

```

Qy      260 YLKKGKBPDPDFDRSHLKLFYIKYVDVDTNELLKSEQLLTASERNLDFRDIYDPRDK 317
       |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db      391 YLRKRIPRPADFVEGGYK---FEWVDVDH---FKIDQLLV--RVLEKNDLPSPPDE 440

RESULT 9
US-10-425-114-59217
; Sequence 59217, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425.114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59217
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700266372_FLI.pep
US-10-425-114-59217

```

Query Match	6.2%;	Score 117;	DB 15;	Length 777;
Best Local Similarity	22.4%;	Pred. No. 1.3;		
Matches	70;	Conservative 48;	Mismatches 138;	Indels 56; Gaps 12;
Qy	69	QEOLIANVHNDNDYFEVIDFASDATITDRNGKVYFADKDGSVTLPTOPVQFELLSGHVRV	128	
Db	366	QEEKMLMARSSKDKFEHLQSQYVDLTSENNALKTEIEELKSRLLIELQRTQETIVMVQHVEE	425	
Qy	129	RPYKEPIQ-----NQAKSVDVVEYTVQFTPLNPDDDDFRPGLKDTKLLKTLAI	175	
Db	426	QVQAEKIRLESEAEVSASINOLEKLASLQRIQKLEDSIFAENHKQELLQKILKL	485	
Qy	176	GDITTSQELLAQASIL-NKNHPGYTY-----ERDSSIVTHNDI--FRTILPMQDEPTY	228	
Db	486	ES--DNQELLQMQOSIMEEKSNNAESLHGHTKRDQQVDTLENQINQLRSVL-----	535	
Qy	229	RVKNREQAY--RINKKSGIAEE-INNTDLISEKYVVLKKGKPPDPPDR-----SHLK	278	
Db	536	--DEKEQYLCSVQKXTLEEQKIQVALLSATCKSDAKKQVDLMLEGEKIELSKHLK	593	
Qy	279	LFTIKYVDVDTNELLKSEQLLTASERNLDRFDLDYDPRDKAKLLYNNLDAFIMDYTLTGK	338	
Db	594	ELSLN-KQ-DAINEIRKKYEKIEIITNAE-----KEKAELIREI-----ENKCKNEK	639	
Qy	339	VEONHDDOTNRII	350	
Db	640	ISQNKHDSERYL	651	

[illegible]

RESULT 10

US-10-425-114-66924
; Sequence 66924, Application US/10425114
; Publication No. US20040034888A1

GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 66924

; LENGTH: 798

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: UC-ZMROMO17068D12_FLI.pep

US-10-425-114-66924

Query Match 6.2%; Score 117; DB 15; Length 798;

Best Local Similarity 22.4%; Pred. No. 1.4; Mismatches 138; Indels 56; Gaps 12;

Matches 70; Conservative 48

QY 69 QBLIANVHNSDDYFEVIDFASDATITDRNGKVFADKDGSVTLPTQPVQEFLLSGHVRV 128

DB 259 QEEKMLMARSSKDEHLOSQVDLTSENALKTEELKSLRIELQRTQEIWMVQHVVE 318

QY 129 RPYKEKPIQ-----NQAKSVDEVYVQFTPLNPDDDFRPGKOTKLLKTLAI 175

DB 319 QVAVEDKIRRLSEAEVSNINQLEKASELQGRIOKLELSTFAENHKQELLQKILK 378

QY 176 GDTITSQELLAAQAIL-NKNHPGYTIY-----ERDSSIYTHNDI--FTILPMDQEFY 228

DB 379 ES--DNQELLGQMGEIMEEKSNNASLHGEITKRDQOQVDTLENQINQLRSVL----- 428

QY 229 RVKNEQAY--RINKSGLNEE- INNTDLISEKYVYVKKGERPYDPDR-----SHLK 278

DB 429 --DEKQLVCSQVEKTELEOKLQVEALLSATECKLSAKQYDLMLEGEKIELSKHLK 486

QY 279 LFTIKYVDVTNELKSEQLLFASERNLDFRDLDPDRKAKLLYNNLDAFGIMDYTLTGK 338

DB 487 ELSLKN-DQAINERIKYVELEKIEITNAE-----KEAKELIREI-----ENKCNEK 532

QY 339 VEDNHDNTNRII 350

DB 533 ISQKHDSERYL 544

RESULT 11

US-10-282-122A-75177

; Sequence 75177, Application US/10282122A

; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELTRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 75177

; LENGTH: 5559

; TYPE: PRT

; ORGANISM: Salmonella typhimurium

US-10-282-122A-75177

Query Match 6.1%; Score 115.5; DB 15; Length 5559;

Best Local Similarity 20.3%; Pred. No. 25;

Matches 77; Conservative 49; Mismatches 158; Indels 85; Gaps 13;

QY 17 DISLKEFEIDLSRPAHGKTEOGLSPKSPFATDSCAMSHKLEKADLLKA----- 67

DB 4590 DSQIAVFDIDEDSLPALSNNRALSVSQVGEAGSOVSIFVDGKLVNVMVEADGTWRAPIL 4649

QY 68 IQBOLIANVHNSDDYFEVIDFASDATITDRNGKVFADKDGSVTLPTQPVQEFLLSG--- 124

DB 4650 LQDDGTFTNIH-----FSITDVAGNTEVS-----KDYSVDVDSSTDFTLNLEDASNSGSLD 4700

QY 125 -----HVRVPYKEKPIQNAQSVDEVYVQFTPLNPDDDFRPG 163

DB 4701 DLITNHNKPVLVGTAEAGATIHIV---DEKIVANVLVLEDGTWSYQF-----DNALKOG 4752

QY 164 LKDTLLKTLAIGDTTTSQELLAAQOAILNKNHPGYT-----IYERDSSIYTHNDIFR 217

DB 4753 EYSIRVVAEDPAGNTAESPELLVTIDTSTFIDNPAMVAGSDNGIFNSD-SITSQTRTFSS 4811

QY 218 TILPMDQEF-----TYRVNREQAVRINKKGLNEEINNNTDLISEKYVYL--KKG 265

DB 4812 IFGEMNQSVQIFIDGVLVDITITVTRNQVVRPESPLG-----DGSHSIYVYVITDKAG 4863

QY 266 EKYPDFDRSHLKLFTIKYVDVDT-----NELLKSEQ--LLTASENLDFRDLDPDRK 317

DB 4864 NTATSKTLNFTIDTFTNTTTPVAIDSIGQTLAEMTSGDGKIYITDTTRNLLFSGSAEPNSK 4923

QY 318 AKLLYNNLD 326

DB 4924 IEIINGLN 4932

RESULT 12

US-09-919-703-7

; Sequence 7, Application US/09919703

; Patent No. US20020165129A1

; GENERAL INFORMATION:

; APPLICANT: Krystal, Gerald

; APPLICANT: Rabkin, Simon W.

; TITLE OF INVENTION: Peptides and Their Use to Ameliorate

; TITLE OF INVENTION: Cell Death

FILE REFERENCE: 50216/003004
CURRENT APPLICATION NUMBER: US/09/919,703
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 09/294,457
PRIOR FILING DATE: 1999-04-19
PRIOR APPLICATION NUMBER: US 08/759,599
PRIOR FILING DATE: 1996-12-05
PRIOR APPLICATION NUMBER: US 60/008,233
PRIOR FILING DATE: 1995-12-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic polypeptide
US-09-919-703-7

Query Match 6.0%; Score 115; DB 9; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 SVDVEYTVQFTPLNPDHFRP 162
DB 1 SVDVEYTVQFTPLNPDHFRP 21

RESULT 13

US-10-282-122A-71575
Sequence 71575, Application US/10282122A
Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

PRIOR FILING DATE: 2001-02-09

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614

SOFTWARE: PatentIn version 3.1

SEQ ID NO 71575

LENGTH: 561

TYPE: PRT

ORGANISM: Staphylococcus haemolyticus

US-10-282-122A-71575

Query Match 6.0%; Score 113.5; DB 15; Length 561;

Best Local Similarity 18.7%; Pred. No. 1.6;

Matches 74; Conservative 68; Mismatches 147; Indels 107; Gaps 17;

QY 8 AGTVEGTNODISLKFFPEI-----DLTSPAHGGKTEQGLSPKSKPFATDSGAMSH 57

DB 52 SANIESTNKPLSLESSNVSHNEQATPSSETIHPSTVETSTTQNTTSKHNSNQSQAFA 110

QY 58 KLEKADLLKAIQELIANVHNSDYFEVIDFASDATITDRNGKYF-----ADKDG 108

DB 111 -----QEKATPEPATNDNTSSQEKASSESTSNENNTSQFHGTEVKSPSPSEQ 159

QY 109 SVTLPTQPVQEFLLSGHVRVPYKEPIQNAQKSVDEYTVQFTPLNPDHFRPGLKDTK 168

DB 160 SKSTQVEPTKE---STSTKVQTPQBPQOQKR-NVPESQQSTPAS-----202

QY 169 LKATLAIGDITTSOELLAQAQSIILKNHPCYTIYERDSSIVTHDNDI-----215

DB 203 --KQVSKSSNTTSEQAAPKSOHVTVNKRMAATTSDNTNKTILHTNDIHGRFVDDGRVIG 260

QY 216 -----FRTILPMDQETFR---VKNREQAYRINK-KSGL-----NEEIN---250

DB 261 MAKVKGKDKYNPDLMVDSGDAFQGLPVSNNKGEEMAKAMNGVGYDAMTVGNHFDG 320

QY 251 -----NTDLISEKYVYLKKGKPYDPDRSHLKLFTIKY--VDVDTNELLKSEQL 298

DB 321 DQLLKLQKQLNFPFIVSSNIY--KNGKRVFDP--STTVTKNNVRYGIVGVTTPE-TTKTKS 375

QY 299 LTRASERNLDPD-LYDPRDKAKLLYNLDAGIMDY 333

DB 376 PTAVE-GVEFKDPLTSVKQAMNEIKNNVDVFVILSH 410

RESULT 14

US-10-282-122A-76865

Sequence 76865, Application US/10282122A

Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

this is a NIH_MGC Library."

```
ORIGIN
Alignment Scores:      0.0427      Length:      948
Pred. No.:            109.00      Matches:      60
Score:                42.66%      Conservative: 33
Percent Similarity:   27.52%      Mismatches:  71
Best Local Similarity: 5.73%      Indels:      54
Query Match:         4           Gaps:        12
DB:

US-09-940-235-2_COPY_16_383 (1-368) x B1522500 (1-948)
Qy 29 SerArgProAlaHisGlyLysThrGluGlnGlyLeuSer-ProLysSerIysProph 48
Db 365 TCTGCTCCCTACACAGGGGAAAACAGGAGAAAGTCTGGTTCTTAAGCAGAGCCACT 424
Qy 48 e-----AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 66
Db 425 GGATCCTGTGCTGCTGAAAGTGTGACGCTGGAACCGGAGCTGGAGGAAGCC---TTGGCAA 481
Qy 66 sAlaIleGlnGluGlnLeu-----lleAlaAsnValHis-----Se 78
Db 482 TGCCTTCAGATGCAGAACTCTGTGACATTCGACGATCTGGGGCATGCACAGCTCATGAG 541
Qy 78 rAsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAs 98
Db 542 TAACACGACGACTACTACCGCCCTG-----AGCAGCAGCTCCATCATGAACAAGGA 592
Qy 98 nGlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValG 118
Db 593 GGGGCTC-----AACAGCGTGTATTAACCCACACAA-----623
Qy 118 nGluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAs 138
Db 624 -----TACAAGCCGTGTC-----GACGA 643
Qy 138 nGlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAs 158
Db 644 AGAACCAAAATTCACAGAGCTAGAGGAACGCTG-----GA 679
Qy 158 pAspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspTh 178
Db 680 ACGGATAAGAACACACGACCCAAACTTGAAGAAGTTAACTCAATATATATCCGA-----734
Qy 178 rIleThrSerGlnGluLeuAlaGlnAlaGlnSerIleLeuLysAsnLysAsn-HisProG 198
Db 735 -ATATCCCATCCCATCCCTCAAGGCATTTCAGAAAGCC---CTGAAGAAAAAATCAT---786
Qy 198 lyTyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgT 218
Db 787 -----ATGTCAGACAAAGTTTCAGATCCGTCGGGAGCCGGAGTAATGACCCGTCGGTA 838
Qy 218 hrIleLeuProMetAspGlnGluPheThrTyrArgValLysAsnArgGlu 234
Db 839 ATGCCCTCTCTCGGATGCTCAGGAGAACCAAGTGTTCAGACACACTGAA 888

RESULT 3
CNS07BAJ/c      1128 bp      DNA      linear      GSS 08-JUL-2001
LOCUS          T7 end of clone BC0AA010E09 of library BC0AA from strain CBS 767 of
DEFINITION      Debaryomyces hansenii, genomic survey sequence.
ACCESSION      AL437633
VERSION        AL437633.1  GI:12221046
KEYWORDS       GSS.
SOURCE         Debaryomyces hansenii (anamorph: Candida famata)
ORGANISM       Debaryomyces hansenii
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Debaryomyces.
1 (bases 1 to 1128)
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
PDBS Lett. 487 (1), 3-12 (2000)
20584711
PUBMED 11152876
REFERENCE 2 (bases 1 to 1128)
Leplingle,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H.,
Artiguenave,F., Wincker,P. and Gaillardin,C.
Genomic exploration of the hemiascomycetous yeasts: 14.
Debaryomyces hansenii var. hansenii
PDBS Lett. 487 (1), 82-86 (2000)
20584724
PUBMED 11152889
REFERENCE 3 (bases 1 to 1128)
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
segrf@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.

FEATURES
Location/Qualifiers
source          1..1128
                /organism="Debaryomyces hansenii"
                /mol_type="genomic DNA"
                /strain="CBS 767"
                /variety="hansenii"
                /db_xref="taxon:4959"
                /clone="BC0AA010E09"
                /clone_lib="BC0AA"
                /notes="end : T7"
                /notes="end : >1030"
                /notes="similar to Saccharomyces cerevisiae ORF YIL005w [
                similarity to protein disulfide isomerases ]"
                /evidence=not_experimental

misc_feature
                complement (<5..>1030)
                /notes="similar to Saccharomyces cerevisiae ORF YIL005w [
                similarity to protein disulfide isomerases ]"
                /evidence=not_experimental

ORIGIN
Alignment Scores:      0.0548      Length:      1128
Pred. No.:            109.00      Matches:      60
Score:                39.38%      Conservative: 42
Percent Similarity:   23.17%      Mismatches:  81
Best Local Similarity: 5.73%      Indels:      76
Query Match:         9           Gaps:        14
DB:

US-09-940-235-2_COPY_16_383 (1-368) x CNS07BAJ (1-1128)
Qy 51 AspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLysAlaIleGlnGlu 70
Db 721 GACTTGAAGCTTTTTCGACGTCGGTTATACGAGAGTTCTCAATATGAAAGATATCTGCT 662
Qy 71 GlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIleAspPheAlaSer 90
Db 661 GCCACGTTATCGGAGGTAATG-----GAGTATAGAAATAATTTACCTTTTGACACCT 611
Qy 91 AspAlaThrIleThrAspArgAsnGlyLys-----ValTyrPheAlaAspLysAsp 107
Db 610 TTGAATCTTATTATATCTCTTTTAATAATAAGATATACGATGCTCTTTTATACGATGAA 551
Qy 108 GlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArg 127
Db 550 ---ACGGTAACAGAGGATAGGCCATTTTACCATTATTTATTTGGAG---TATGTAATG 497
```



```

Db 173 GTCAGGCTTAAGTCCACTCCCGCCCTAAGTTC-----TCTGTGTGT 214
Qy 130 ProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyr----- 147
Db 215 GTCCTGGGGGACAGCAGCAGTGTGACAGGCTAGCGGTAGCGTGGATATCCCCACATGGAC 274
Qy 148 ThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThr 167
Db 275 ATCGAGGCGCTCAAAAACCTCAAC-----AAGAATAAA 307
Qy 168 LysLeuLeuLysThrLeuAla-----IleGlyAspThrIleThrSerGlnGluLeuLeu 185
Db 308 AAATCGGTCAAGAAGCTGCCAAGAAGTATGATGCGCTTTTGGCCCTCAGAGTCTCTGATC 367
Qy 186 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 205
Db 368 AAGCAGATCCACGAATCTCTCGC-----CCAGGTTTAAATAGGAGGAGAAAGTTC 418
Qy 206 SerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGlu 225
Db 419 CTTCTCCCTGCTCACACACAGAAACATGGTGGCCAAAGTGGATGAGTGAAGTCCACA 478
Qy 226 PheThrTyrArgValLys----- 231
Db 479 ATCAAGTTTCCAATGAAGAAGGTGTTATGCTGCTGTAGTCTGTTGTCACGTGAAGATG 538
Qy 232 ---AsnArgGluGlnAlaTyrArgIle----- 239
Db 539 ACAGACGATGAGCTGTGTATTAACATTTCACTGGCGCTGTCAACTTCTTGGTGTCAATGC 598
Qy 240 AsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyr 259
Db 599 TCAAGAAAAAAGTGGC-----AGAATTGTCGGGCTTATAT 634
Qy 260 TyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeu 279
Db 635 ATCAAGAGCANCATGGCAAGCCCGCCAGCGCCCTATATTAAAGGCCCATTTTGAATTA 694
Qy 280 PheThrIleLysTyrVal 285
Db 695 AATTCTATTACCAAGTT 712

RESULT 10
BX861601/c
LOCUS
DEFINITION BX861601 tcba Oncorhynchus mykiss cDNA clone tcba0014c.o.16 5prim,
mRNA sequence.
ACCESSION BX861601
VERSION BX861601.2 GI:42758263
KEYWORDS EST.
SOURCE Oncorhynchus mykiss (rainbow trout)
ORGANISM Oncorhynchus mykiss
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
1 (bases 1 to 721)
Govoroun, M., Guiguen, Y. and Le Gac, F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished (2003)
On Dec 16, 2003 this sequence version replaced gi:39959737.
Contact: Guiguen Y
INRA - SCRIBE
Campus de Beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signena@portejou.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0014 row: 0 column: 16
Seq primer: M13R.

```

```

FEATURES
source
Location/Qualifiers
1..721
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcba0014c.o.16"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="tcba"
/notes="Vector: pT73D-pac; Rainbow trout multi-tissues -
normalized + 1 subfraction; Clone distribution: AGENAE
Resource centre, Francois PIUMI,
Francois.PIUMI@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33
(0) 1.34.65.22.73"

ORIGIN
Alignment Scores:
Pred. No.: 0.116 Length: 721
Score: 104.00 Matches: 51
Percent Similarity: 32.41% Conservative: 43
Best Local Similarity: 17.59% Mismatches: 82
Query Match: 5.47% Indels: 114
DB: 5 Gaps: 11

US-09-940-235-2_COPY_16_383 (1-368) x BX861601 (1-721)

Qy 106 LysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHis 125
Db 680 AAGCCCGGGTGATGTCGACTCTCTGTTAAAGGTTCCCGAGGAGCCCTAATCCAGGCC-- 624
Qy 126 ValArgValArgProTyr-----LysGluLysPro 135
Db 623 ATCAGCGTCCAGCCACTGGCCAGAGAGGACAAACTGCCCGCTGCCAGAACCAAGCCCA 564
Qy 136 IleGlnAsnGlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsn 155
Db 563 GTGTATGCCAGCTGGCCAGCTGATGTGGAC-----CTGCCAGTCAGC 519
Qy 156 ProAspAspAsp-----PheArgProGlyLeu 164
Db 518 CCCTCTGACGCCCGCAGTCCCGCTGGTCTATGATGACAGCATCTCCAGGCCGAGTATG 459
Qy 165 LysAspThrLysLeuLeuLysThrLeuAlaIleGly----- 176
Db 458 AAGCTGTTGAAGTTTAAAGAGGGGAGAGTGTGGCTGAGGCTGGCCGAGGAACGAC 399
Qy 177 -----AspThrIleThrSerGlnGluLeuAla 186
Db 398 GTGGGCATCTTTAGCCGGAGTCTGGAGGAGCAGCCAGCTGCCAAGGAAGTCTGGAG 339
Qy 187 GlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAspSer 206
Db 338 GAAGGAGACAGATTCTCAGGTTAATAACGTAGACTTT-----GCT 297
Qy 207 SerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGlnGluPhe 226
Db 296 AACATCATCCGAGAGGAGGCTGTGCTCTCTAGACCTTCCAGGGGTGAAGAGGTC 237
Qy 227 ThrTyrArgValLysAsnArgGluGlnAlaTyrArgIleAsnLysLysSerGlyLeuAsn 246
Db 236 ACCATTCTGCCCCAGAGAGAGAGAGTGTTCCTGCT----- 201
Qy 247 GluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysGlyGlu 266
Db 200 ---CGATAGTAGGTCGGAT---GTGGGCGACTCGTCTTACATC----- 162
Qy 267 LysProTyrAspProPheArgSerHisLeuLysLeuPheThrIleLysTyrValAsp 286
Db 161 -----CGAACCCACTTTGAG----- 147

```


Qy 287 ValAspThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArgAsnLeu 306
Db 147 ----- 147

Qy 307 AspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsnLeuAsp 326
Db 146 ----- 146

Qy 327 AlaPheGlyIleMetAspTyrThrLeuThrGlyLys----- 338

Db 101 GTGTTCCGTGAGTGGGACCCCTTACACAGCAAGCTGGCTCTGCTGGCTGCCATACGC 42

Qy 339 ValGluAspAsnHisAspThrAsnArg 348
Db 41 ATTGGCAAGAACCCAGAGAGTGGAGAGA 12

RESULT 11
AW640519 765 bp mRNA linear EST 26-APR-2001
LOCUS b196a08.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
DEFINITION laevis cDNA clone PX0096A08 5', mRNA sequence.

ACCESSION AW640519
VERSION AW640519.1 GI:7397747
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE 1 (bases 1 to 765)
AUTHORS Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.
Jr., Moore, D.T., Bonfield, G.G., Beckstrom-Sternberg, S.M.,
Touchman, J.W., Bonaldo, M.F., and Soares, M.B.
The NIEHS Xenopus maternal EST project: interim analysis of the
first 13,879 ESTs from unfertilized eggs
Gene 267 (1), 71-87 (2001)

JOURNAL 21211403
MEDLINE 11311557
PUBMED

COMMENT Contact: Perry J. Blackshear
Office of Clinical Research and Laboratory of Signal Transduction
National Institute of Environmental Health Sciences
A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
USA

Tel: 919 541-4899
Fax: 919 541-4571

Email: black009@niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial
Parkway, Huntsville, AL 35901

phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
cdna@resgen.com

DNA Sequencing and analyses performed by National Institutes of
Health Intramural Sequencing Center (NISC).

PCR Primers

FORWARD: TGTAACGACGGCCAGT

BACKWARD: CAGAACAGCTATGACC

Seq primer: 0096 row: A column: 08

Location/Qualifiers

1..765

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="PBX0096A08"

/sex="female"

/tissue_type="unfertilized egg"

/cell_type="unfertilized egg"

/dev_stage="unfertilized egg"

/lab_host="DH10B"

/clone_lib="Blackshear/Soares normalized Xenopus egg
library"

/note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI;

PolyA-selected mRNA was prepared from unfertilized Xenopus

laevis eggs. The library was constructed in the vector

pT73-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalize and subtract: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dT18 primer; double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT73-Pac vector.
The library contained approximately 7.2 x 10⁵
recombinants, with average insert sizes of 1-1.5 kb."

ORIGIN

Alignment Scores:
Pred. No.: 0.126 Length: 765
Score: 104.00 Matches: 51
Percent Similarity: 43.24% Conservative: 45
Best Local Similarity: 22.97% Mismatches: 86
Query Match: 5.47% Indels: 40
DB: 2 Gaps: 11

US-09-940-235-2_COPY_16_383 (1-368) x AW640519 (1-765)

Qy 47 PropheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeuLys 66

Db 88 CCAATTGTCAATGAGTACACCAAGCTCTGAGTCAATAATGTTTCGACGCTGCT 147

Qy 67 AlaIleGlnGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIle 86

Db 148 GCTGAGAAGAGACGACACACAACTCTCTGAGACTTCAACGCTCAACAAGCCACTT 207

Qy 87 AspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyr-----Phe 103

Db 208 GGCCTTACTGTTGATAATCAAGAAGTTGAAGGGGATGGAGAACCTTATGCAACAGAGTTT 267

Qy 104 AlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSer 123

Db 268 ATAGACAGCTCAAGCAGTTAGTGTTCCTCTCT-----TCC 309

Qy 124 GlyHisValArgVal-----ArgProTyr-----LysGluLysProIleGlnAsn 138

Db 310 AGCCAGGTTAAATATGAGACTGCAGAGCCTTATCTTGGTGATAGAAATATCGGAAT 369

Qy 139 GlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAsp 158

Db 370 GAGAAAAAACCGAGACCTTGAGCAGAAAAAGTAAG-----AATCCAGAGAA 417

Qy 159 AspPheArgProGlyLeuLys-----AspThrLysLeuLeuLysThrLeuAla 174

Db 418 GACGATTATCCCAAGTCAAGTTAAATCTCCAGAAAGATCAATACTGGAGGAACTGTT 477

Qy 175 IleGlyAspThrIleThrSerGlnGluLeuAlaGlnSerIleLeuLeuLys 194

Db 478 -----AAAGAGATAGTTGCCAAGGATAAATCTCTGCTGAATGTC 516

Qy 195 AsnHisProGlyTyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAsp 214

Db 517 CTCACGCT-----GTTTCAGTCAGAAATCTGCTATG-----GAT 552

Qy 215 IlePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgValLysAsnArgGlu 234

Db 553 CTCATGAAAAGCCTCTTCCTATGGAC-----GTCACTGCAGCGGAAAAAATCTAGA 606

Qy 235 GlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGluGluLeuLeuLeuLeuLeu 254

Db 607 AGAGGACTT---CTTGGAAAAAGCAAGGGGAAACATTAAAGAAAAACAACCTCGG 663

Qy 255 IleSer 256

Db 664 GAAAGT 669

RESULT 12

AW640593 765 bp mRNA linear EST 26-APR-2001

LOCUS b196h01.w1 Blackshear/Soares normalized Xenopus egg library Xenopus

DEFINITION

laevis cDNA clone PBX0096H01 5', mRNA sequence.

ACCESSION AW640593
 VERSION AW640593.1 GI:7397825
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 Xenopus laevis; Xenopus.
 REFERENCE 1 (bases 1 to 765)
 AUTHORS Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.
 Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M.,
 Touchman, J.W., Bonaldo, M.F., and Soares, M.B.
 The NIHES Xenopus maternal EST project: Interim analysis of the
 first 13,879 ESTs from unfertilized eggs
 JOURNAL Gene 267 (1), 71-87 (2001)
 MEDLINE 21211403
 PUBMED 11311557
 COMMENT Contact: Perry J. Blackshear
 Office of Clinical Research and Laboratory of Signal Transduction
 National Institute of Environmental Health Sciences
 A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
 USA
 Tel: 919 541-4899
 Fax: 919 541-4571
 Email: black009@niehs.nih.gov
 Clone is available through Research Genetics, Inc., 2130 Memorial
 Parkway, Huntsville, AL 35901
 phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
 cdna@resgen.com
 DNA Sequencing and analyses performed by National Institutes of
 Health Intramural Sequencing Center (NISC).
 PCR Primers
 FORWARD: TGTAAACGACGCCAGT
 BACKWARD: CAGGAAACAGTATGACC
 Plate: 0096 row: H column: 01
 Seq primer: T7 primer.

FEATURES

Location/Qualifiers
 1..765
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="PBX0096H01"
 /sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /clone_lib="Blackshear/Soares normalized Xenopus egg
 library"
 /note="Vector: pT7T3-Pac; Site 1: EcoRI; Site 2: NotI;
 PolyA-selected mRNA was prepared from unfertilized Xenopus
 laevis eggs. The library was constructed in the vector
 pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
 Soares, M.B. 'Normalization and subtraction: two
 approaches to facilitate gene discovery', Genome Research
 6:791-805, 1996. The first strand synthesis used a
 NotI-dt18 primer; double stranded cDNAs were ligated to
 EcoRI adapters, digested with NotI, and directionally
 cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
 The library contained approximately 7.2 X 10⁵
 recombinants, with average insert sizes of 1-1.5 kb."

ORIGIN

Alignment Scores:
 Pred. No.: 0.126 Length: 765
 Score: 104.00 Matches: 51
 Percent Similarity: 43.24% Conservative: 45
 Best Local Similarity: 22.97% Mismatches: 86
 Query Match: 5.47% Indels: 40
 DB: 2 Gaps: 11

US-09-940-235-2_COPY_16_383 (1-368) x AW640593 (1-765)

Qy 47 ProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLys 66
 ||| : : : : :
 Db 88 CCAATTGTCAGTGAGACTACACCAAGCTCTGAGTCTGATGAAATTTTCGACGCTGCT 147
 ||| : : : : :
 Qy 67 AlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGluValIle 86
 ||| : : : : :
 Db 148 GCTGAGAAAGAGACGACGACACCAACCAACTTCTCTGAGACTTCCAAACGTCACCAAGCCACTT 207
 ||| : : : : :
 Qy 87 AspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyr-----Phe 103
 ||| : : : : :
 Db 208 GGCTTTACTGTTGATAATCAAGAAGTTGAAGGGATGGAACCTATGCAACACAGAGTTT 267
 ||| : : : : :
 Qy 104 AlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLys 123
 ||| : : : : :
 Db 268 ATAGACAGCTCAAGACGATTAGAGTTGAGTTCCTCTCT-----TCC 309
 ||| : : : : :
 Qy 124 GlyHisValArgVal-----ArgProTyr-----LysGluLysProIleGlnAsn 138
 ||| : : : : :
 Db 310 AGCCAGGTTTAAACATAATGCAGACTGCAGAGCTTATCTTGGTGATAAGAAATATCGGAAT 369
 ||| : : : : :
 Qy 139 GlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAsp 158
 ||| : : : : :
 Db 370 GAGAAAAAAGCGAAGACCTTGACGAGAAAGTAAG-----AATCCAGAGAA 417
 ||| : : : : :
 Qy 159 AspPheArgProGlyLeuLys-----AspThrLysLeuLysThrLeuAla 174
 ||| : : : : :
 Db 418 GACGATTATCCCAAGGTCAGTTAAATCTCCAGAGATGAATACTGGAGAACTGTT 477
 ||| : : : : :
 Qy 175 IleGlyAspThrIleThrSerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLys 194
 ||| : : : : :
 Db 478 -----AAAGAGATAGTTGCCAAGGATAAATCTCTGCTGAATGTC 516
 ||| : : : : :
 Qy 195 AsnHisProGlyTyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsn 214
 ||| : : : : :
 Db 517 CTCAGCCT-----GTTTCAGTCAGAGAAATCTGCTATG-----GAT 552
 ||| : : : : :
 Qy 215 IlePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgValLysAsnArgGlu 234
 ||| : : : : :
 Db 553 CTCATGAAAGCCCTCTTCCTATGGAC-----GTCACTGCAGCGGAAATCTAGAACT 606
 ||| : : : : :
 Qy 235 GlnAlaTyrArgIleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeu 254
 ||| : : : : :
 Db 607 AGAGAGCTT---CTTGAAAAAGACAAGGGGGAACATTAAGAAAAACAACCTCGGACCTT 663
 ||| : : : : :
 Qy 255 IleSer 256
 ||| : : : : :
 Db 664 GAAAGT 669

RESULT 13

BE393232
 LOCUS BE393232
 DEFINITION 601308437F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3626601 5',
 mRNA sequence.
 ACCESSION BE393232
 VERSION BE393232.1 GI:9338597
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 657)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLCM308 row: j column: 10
High quality sequence stop: 604.
Location/Qualifiers

FEATURES
source

1..657
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3626601"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 44"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred. No.: 0.117 Length: 657
Score: 103.50 Matches: 59
Percent Similarity: 39.74% Conservative: 32
Best Local Similarity: 25.76% Mismatches: 74
Query Match: 5.44% Indels: 64
DB: 2 Gaps: 14

US-09-940-235-2_COPY_16_383 (1-368) x BE393232 (1-657)

Qy 29 SerArgProAlaHisGlyGlyValThrGluGlnGlyLeuSer-ProLysSerLysProPh 48
Db 30 TCTGTCCTCCCTACACAGGGAAGAAAGTCTGGTCTCTTAAGAGCAAGCACT 89
Qy 48 e-----AlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAspLeuLeu 66
Db 90 GGATCCTGTGCGAAGTGTACGCTGGAAACCGGAGCTGGAGAGCC---TTGGCAAA 146
Qy 66 sAlaIleGlnGluGlnLeu-----lleAlaAsnValHis-----Se 78
Db 147 TGCTTCAGATGCAAGCACTCTGTGACATTCAGCGATCTCTGGGATGCACACGCTCATGAG 206
Qy 78 rAsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAs 98
Db 207 TAACCAAGCAGTACTACACAGGCTG-----AGCAGCAGCTCCATCATGAACAAGA 257
Qy 98 nGlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGl 118
Db 258 GGGGCTC-----AACAGCGTGATTAACCCACACAA----- 288
Qy 118 nGluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAs 138
Db 289 -----TACAGCCTGTGCC-----CACGA 308
Qy 138 nGlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAs 158
Db 309 AGAACCAAAATTCACACAGAGCTAGAGGAAACGCTGGAAACGATAAAGAACACACAC 363
Qy 158 pAspPheArgProGlyLeuLysAspThrLysLeu-----LeuLysThrLeuAlaIleGl 176
Db 364 -----CCAAACTTGAGAGAGTTAACTCAATATATATCCGGAATATCCCCATC-- 411
Qy 176 yAspThrIleThrSerGlnGluLeuAlaGlnSerIleLeuAsnLysAsnHi 196
Db 412 -----CCACACCTCAAGGCATATGCGAAGCCCTGAAGAAACCTCA-- 453
Qy 196 sProGlyTyrThrIleTyrGluArgAspSerIleVal-----ThrHisAspAsnAspI 215
Db 454 -----TATGTGAAGAGAGTTTCAGCATCGTGGGACACGCGAGTAAATGAC-- 495
Qy 215 ePheArgThrIleLeuProMetAspGlnGluPheThrTyrArgValLysAsnArgGluGl 235

Db 496 -----CCCGTGGCGTATGCCCTTGTCTGAGATGCTCAAGGAGAACACGAGT 539
Qy 235 nAlaTyrArgIleAsnLysLysSer 243
Db 540 GTTGAAGACACTGAATGTGGATCC 564

RESULT 14

CD556206
LOCUS CD556206 1058 bp mRNA linear EST 11-JUN-2003
DEFINITION AGENCOURT_14401188 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CD556206
VERSION CD556206.1 GI:31582274
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1058)
AUTHORS NIH-MGC http://img.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDKM69 row: d column: 23
High quality sequence start: 5
High quality sequence stop: 582.

FEATURES

source

1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic trophoblasts, made from WA01 stem cells"
/lab_host="DH10B Tona"
/clone_lib="NIH_MGC_173"
/note="Vector: pDONR201; Site 1: attP2; Site 2: attP1;
LIBR PRIMING - oligo dt; METHOD - full-length enriched;
LIBR PROVIDER - Bradfield"

ORIGIN

Alignment Scores:
Pred. No.: 0.232 Length: 1058
Score: 103.50 Matches: 63
Percent Similarity: 38.43% Conservative: 45
Best Local Similarity: 22.42% Mismatches: 95
Query Match: 5.44% Indels: 79
DB: 6 Gaps: 11

US-09-940-235-2_COPY_16_383 (1-368) x CD556206 (1-1058)

Qy 44 LysSerLysProPheAlaThrAspSerGlyAlaMetSerHisLysLeuGluLysAlaAsp 63
Db 15 AAATCTTTTCGGTTAGCGCGCGTGAAGAGCCATGACGACGAAGTCTCTCGGACACC 74
Qy 64 LeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspTyrPhe 83
Db 75 CTGTACAGGCGGTGCGGAGTCTCTGCACGGGAACACGACGCGCCCAAGTACCTG 134
Qy 84 GluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPhe 103
Db 135 GAGACGCTGGAGATGACATGACATGCTTG-----AAGAATCATGAT 173
Qy 104 AlaAspLysAsp-----GlySerValThrLeuProThrGlnProValGlnGlu 119

```
Db 174 CCCCAGAGCAAGCGCTTCGGGCACCGTTCAGGCTTAAGTCCATCTCCGCCCTTAAG 233
Qy 120 PheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGln 139
Db 234 TTC-----TCTGTGTGTCTCTGGGGACGACGAGCACTGTGAGGAG 275
Qy 140 AlalysSerValAspValGluTyr-----ThrValGlnPheThrProLeuAsnProAsp 157
Db 276 GCTAAGGCGGTGGATATCCCCACATGGCATCGAGCGGTGAAAAAATCAAC----- 329
Qy 158 AspAspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAla-----Ile 175
Db 330 -----AAGATAAAAACTGTCTCAAGAGGTGGCCCAAGAGTAT 368
Qy 176 GlyAspThrIleThrSerGlnGluLeuAlaGlnAlaGlnSerIleLeuAsnLysAsn 195
Db 369 GATGGTGTTCGCTCAGTCTCTGATCAAGCAGATCCACGATCTCTCGGC----- 422
Qy 196 HisProGlyTyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIle 215
Db 423 ---CCAGGTTTAAATAAGCAGGAAAGTTCCTTCTCTGCTCACACACAGCAAAACATG 479
Qy 216 PheArgThrIleLeuProMetAspGlnGluPheThrTyrArgValLys----- 231
Db 480 GTGGCCAAAGTGGATGAGGTGAAGTCCCAATCAAGTTCCAATGAAGAGGTGTATGT 539
Qy 232 ------AsnArgGluGlnAlaTyrArgIleAsn 240
Db 540 CTGGCTGTAGCTGTGGTGCACGTGAAGATGACAGCAGATGAGCTGTGTATAACATTCC 599
Qy 241 -----LysLysSerGlyLeuAsn----- 246
Db 600 CTGGCTGTCAACTTCTGTGTGTCTCTCAAGAAACTGGCAG-AATGTCCGGGCCCTT 658
Qy 247 -----GluGluIleAsnAsnThrAspLeuIleSer 256
Db 659 ATATATCAAGAACCACTGGGCAAAACCCAGCGCCCTATATTAAAGGACATTTGAAT 718
Qy 257 GluLysTyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHis 276
Db 719 AAATCTTATTACAGTTAAATAAAAAA-----AAAAACCCAC 760
Qy 277 Leu 277
Db 761 CTT 763

RESULT 15
BOQ59207
LOCUS BOQ59207
DEFINITION AGENCOURT_7049866 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5815208
5', mRNA sequence.
ACCESSION BOQ59207
VERSION BOQ59207.1 GI:19818547
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1073)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2070 row: b column: 09
```

```
High quality sequence stop: 591.
Location/Qualifiers
source 1..1073
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
clone="IMAGE:5815208"
tissue_type="lymphoma, cell line"
lab_host="DH10B (phage-resistant)"
clone_lib="NIH_MGC_99"
notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 0.237 Length: 1073
Score: 103.50 Matches: 62
Percent Similarity: 36.13% Conservative: 37
Best Local Similarity: 22.63% Mismatches: 84
Query Match: 5.44% Indels: 91
DB: 5 Gaps: 12

US-09-940-235-2_COPY_16_383 (1-368) x BQ059207 (1-1073)
Qy 54 AlaMetSerHisLysLeuGluLysAlaAspLeuLysAlaIleGlnGlnLeuIle 73
Db 14 GCATGAGCAGCAAAAGTCTCTCGGCACACCTGTACGAGCGGTGGGAAAGTCTGCAC 73
Qy 74 AlaAsnValHisSerAsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThr 93
Db 74 GGGAAACAGCGCAGCGCCGCAAGTCTCTGGAGACGGTGGAGTTGCAGATCAGCTTG-- 130
Qy 94 IleThrAspArgAsnGlyLysValTyrPheAlaAspLysAsp-----GlySer 109
Db 131 -----AAGAACTATGATCCCCCAGAGGACAGCGCTCTTCGGGGACCC 172
Qy 110 ValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGlyHisValArgValArg 129
Db 173 GTCAGGCTTAAGTCCACTCCCGCCCTAAGTTC-----TCTGTGTGT 214
Qy 130 ProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAspValGluTyr----- 147
Db 215 GTCCTGGGGACACGACGACACTGTGACGAGCTAAGCGCGTGGATATCCCCACATGGAC 274
Qy 148 ThrValGlnPheThrProLeuAsnProAspAspPheArgProGlyLeuLysAspThr 167
Db 275 ATCGAGCGCTGAAAAAACTCAAC-----AAGAATAAA 307
Qy 168 LysLeuLeuLysThrLeuAla-----IleGlyAspThrIleThrSerGlnGluLeu 185
Db 308 AAATGCTGCAAGAGTGGCCCAAGATATGATCGCTTTTGGCTCTCAGAGTCTCTGATC 367
Qy 186 AlaGlnAlaGlnSerIleLeuAsnLysAsnHisProGlyTyrThrIleTyrGluArgAsp 205
Db 368 AAGCAGATTCCAGATCTCTCGGC-----CCAGTTTAAATAAGGACGAGAAATTC 418
Qy 206 SerSerIleValThrHisAspAsnAspIle----- 215
Db 419 CTTTCCCTGTCTACACACACAAAGAAACATGCTGGCCAAAGTGGATGAGTGAAGTCCACA 478
Qy 216 -----PheArgThrIleLeuProMet----- 222
Db 479 ATCAAGTTCCAAATGAAGAGGTGTATGCTGTGCTGTAGCTGTGTGTCACGTGAAGATG 538
Qy 223 ---AspGlnGluPheThrTyrArgVal----- 230
```

```

Db 539 ACAGACGATGAGCTTGTGTATTAACATTCACTGGGTGTCAACTTCTTGGTGTCAATTGCTC 598
Qy 231 ---LysAsnArgGlu-----GlnAlaTyr 237
    |||||
Db 599 AAGAAAAAAGCTGGCAGAAATGTCGGGGCTTATATATCAAGAGCACCATTGGGGCAAGCCCC 658
Qy 238 ArgIleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlu 257
    |||||
Db 659 AGGGCTATATTAAAGCACATTGTAAT-----AAA 688
Qy 258 LysTyrTyrValLeuLysLysGlyGluLysProTyrAspPro 271
    |||||
Db 689 TTCTATTACCACTTAAAGAGAGGAGAAAGACGTAAGACCT 730
    |||||

```

Search completed: November 6, 2004, 04:55:27
 Job time : 4509.03 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:33:59 ; Search time 51.5479 Seconds
(without alignments)
737.669 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQQMVQPSPVAVSQSKPG.....SMIWDCTCIGAGRGISCTI 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	100.0	259	3 AAY90281	Aay90281 Human fib
2	600	100.0	720	2 AAY28914	Aay28914 Fibrinect
3	600	100.0	2324	2 AAR92778	Aar92778 Human fib
4	600	100.0	2324	5 AAU74674	Aau74674 Human fib
5	600	100.0	2324	1 AAE23651	Aae23651 Human pro
6	600	100.0	2327	1 AAP70373	Aap70373 Human fib
7	600	100.0	2327	2 AAR15468	Aar15468 Human fib
8	600	100.0	2328	4 AAG68182	Aag68182 Fibrinect
9	600	100.0	2328	6 ABU07486	Abu07486 Protein d
10	600	100.0	2328	6 ABR41106	Abri41106 Human fib
11	600	100.0	2328	6 ABR92078	Abri92078 Human cer
12	600	100.0	2328	7 ADB70378	Adb70378 Fibrinect
13	600	100.0	2328	7 ADB98726	Adb98726 Human fib
14	600	100.0	2328	7 ADE82522	Ade82522 Human pro
15	600	100.0	2328	8 ADJ37157	Adj37157 Human mal
16	600	100.0	2386	2 AAW63171	Aaw63171 Amino aci
17	600	100.0	2446	2 AAR60021	Aar60021 Fibrinoge
18	600	100.0	2474	4 ABG22279	Abg22279 Novel hum
19	600	100.0	2477	2 AAR95955	Aar95955 Human fib
20	597	99.5	1179	8 ADP75952	Adp75952 Human min
21	597	99.5	1359	8 ADP75957	Adp75957 Human leu
22	596	99.3	231	3 AAB58210	Aab58210 Lung canc
23	596	99.3	463	6 ABR58303	Abri58303 BC00770 p
24	596	99.3	660	2 AAY28901	Aay28901 Human mig
25	596	99.3	1173	4 ABG22275	Abg22275 Novel hum

ALIGNMENTS

RESULT 1

AAAY90281
ID AAY90281 standard; protein; 259 AA.

XX AAY90281;

XX AC

XX 13-OCT-2000 (first entry)

XX DT

XX DE Human fibrinectin protein sequence fragment.

XX KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;

XX KW plasminogen; human; fibrinectin; thrombolytic therapy;

XX KW cardiovascular disorder; fibrinectin.

XX OS Homo sapiens.

XX XX

XX PN EP1024192-A2.

XX XX

XX PD 02-AUG-2000.

XX PF 23-DEC-1999; 99EP-00310541.

XX PR 24-DEC-1998; 98IN-DE0003825.

XX XX

XX PA (COUL) CSIR COUNCIL SCI IND RES.

XX PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;

XX PI Yadav M;

XX XX

XX DR WPI: 2000-516032/47.

XX XX

XX DR N-PSDB; AAA37632.

XX XX

XX PT Hybrid streptokinase-fibrin binding domain polypeptides useful for

XX PT thrombolytic therapy comprises a streptokinase fused with fibrin binding

XX PT domains of human fibrinectin.

XX XX

XX PS Example 3; Fig 6; 58pp; English.

XX XX

XX CC This sequence represents a human fibrinectin fragment, containing fibrin

XX CC binding domains. The invention relates to a hybrid plasminogen activator

XX CC (PA) comprises a polypeptide fusion between streptokinase (SK), which are

XX CC capable of plasminogen (PG) activation, and fibrin binding regions of

XX CC human fibrinectin, which are from fibrin binding domains (RBD) 4 and 5 or

XX CC 1 and 2. The hybrid PA possesses the ability to bind with fibrin

XX CC independently and also characteristically retains a PG activation ability

XX CC which becomes evident only after a pronounced duration, or lag, after

XX CC exposure of the PA to a suitable animal or human PG. The hybrid

XX CC streptokinase-fibrin binding domain polypeptides are useful in

XX CC

XX CC

CC thrombolytic therapy for various kinds of cardiovascular disorders. The
CC hybrids have enhanced fibrin selectivity as well as kinetics of
CC plasminogen activation that are distinct from that of natural
CC streptokinase in being characterised by a temporary delay, or lag of
CC several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase
XX
XX
SQ Sequence 259 AA;

Query Match 100.0%; Score 600; DB 3; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.2e-53;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
DB 1 QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
DB 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106

RESULT 2
AAY28914
ID AAY28914 standard; protein; 720 AA.
XX
AC AAY28914;
XX
XX 21-SEP-1999 (first entry)
XX
XX Fibrinectin protein sequence.
XX
XX Migration stimulatory factor; MSF; cell migration; modulation; human;
XX wound healing; scarring; MSF1-alpha; epitope; fibrinectin.
XX
XX Homo sapiens.
XX
XX WO931233-A1.
XX
XX 24-JUN-1999.
XX
XX 15-DEC-1998; 98WO-GB003766.
XX
XX 16-DEC-1997; 97GB-00026539.
XX
XX (UYDU-) UNIV DUNDEE.
XX
XX Schor SL, Schor AM;
XX
XX WPI; 1999-430039/36.
XX
XX Proteins with cell migration stimulatory activity used in treating wound
XX and preventing scarring.
XX
XX Disclosure; Fig 2; 86pp; English.
XX
XX The invention provides a human migration stimulatory factor (MSF)
XX protein. Host cells containing a replicable vector comprising the MSF
XX encoding nucleic acid can be used for the recombinant production of the
XX protein. The polypeptide can be used for modulating cell migration,
XX healing a wound and for preventing scarring. The present sequence
XX represents the human fibrinectin
XX
XX
SQ Sequence 720 AA;

Query Match 100.0%; Score 600; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 6.8e-53;

Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
DB 50 QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 109
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
DB 110 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 155
RESULT 3
AAR92778
ID AAR92778 standard; protein; 2324 AA.
XX
XX AAR92778;
AC
XX 21-JUN-1996 (first entry)
DT
XX Human fibrinectin.
DE
XX Fibrin-binding peptide; fibrinectin; fibrinolysis; fibrinogenesis;
KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis; therapy.
XX
XX Homo sapiens.
FH
FT Location/Qualifiers
FT 21..65
FT /label= 1F1
FT /note= "type 1 module 1"
FT Domain
FT 66..109
FT /label= 2F1
FT /note= "type 1 module 2"
FT Domain
FT 110..154
FT /label= 3F1
FT /note= "type 1 module 3"
FT Domain
FT 155..199
FT /label= 4F1
FT /note= "type 1 module 4"
FT Domain
FT 200..244
FT /label= 5F1
FT /note= "type 1 module 5"
FT Domain
FT 277..312
FT /label= 6F1
FT /note= "type 1 module 6"
FT Domain
FT 439..486
FT /label= 7F1
FT /note= "type 1 module 7"
FT Disulfide-bond 439..467
FT /note= "Cys439-Cys467 disulfide bond"
FT Disulfide_bind 465..477
FT /note= "Cys465-Cys477 disulfide bond"
FT Domain 487..529
FT /label= 8F1
FT /note= "type 1 module 8"
FT Domain 530..569
FT /label= 9F1
FT /note= "type 1 module 9"
FT Domain 2123..2230
FT /label= Fibrin binding domain
FT /note= "11 kDa C-terminal fibrin-binding domain"
FT Domain 2141..2185
FT /label= 10F1
FT /note= "type 1 module 10"
FT Disulfide-bond 2144..2173
FT /note= "Cys2144-Cys2173 disulfide bond"
FT Disulfide-bond 2171..2183
FT /note= "Cys2171-Cys2183 disulfide bond"
FT Domain 2187..2230
FT /label= 11F1
FT /note= "type 1 module 11"
FT Disulfide-bond 2189..2216
FT /note= "Cys2189-Cys2216 disulfide bond"
FT

XX Modulating proliferation of erythroid progenitor cells in a hematopoietic
PT cellular population by culturing with polypeptides containing an integrin
PT very late antigen 5 or 4 binding site, useful in treating erythroid
PT disorders.

XX Disclosure; Page 55-64; 66pp; English.

XX The present invention relates to a method for enhancing proliferation or
PS inducing apoptosis of erythroid progenitor cells (EPCs). The method
CC involves culturing haematopoietic cellular populations containing EPCs in
CC the presence of a polypeptide having an integrin very late antigen (VLA)-
CC 5 or VLA-4 binding site. The EPCs are useful for treating patients having
CC induced, genetic or congenital erythroid disorders. Methods of the
CC invention are useful in diagnostic assays of progenitor cells, e.g. cells
CC occurring in pathologic states such as anaemia or polycythaemia and in
CC the study of erythroid cell development. They are used in the screening
CC of agents which inhibit or prevent apoptosis of erythroid cells. The
CC present sequence is a human protein related to the modulation of
CC proliferation of EPCs. This sequence is used in the invention

XX SQ Sequence 2324 AA;

Query Match 100.0%; Score 600; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQMVQPSQSPVAVSQSGKPCYDNGKHQIINQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 1 QAAQMVQPSQSPVAVSQSGKPCYDNGKHQIINQWERTYLGVLVCTCYGSGRGFNCESKP 60
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106

RESULT 6

AAP70373
ID AAP70373 standard; protein; 2327 AA.

XX AAP70373;
XX 25-MAR-2003 (revised)
DT 11-MAR-1991 (first entry)

XX Human fibronectin gene product.

XX FN; collagen; fibrin; heparin.

XX Homo sapiens.

XX EP207751-A.

XX 07-JAN-1987.

XX 27-JUN-1986; 86EP-00304998.

XX 28-JUN-1985; 85GB-00016421.

XX (DELZ) DELTA BIOTECHNOLOGY LTD.

XX Baralle FE;

XX WPI; 1987-001441/01.

XX N-PSDB; AAN70596.

XX New fibronectin polypeptide sequence with affinity for collagen etc. -
PT useful for targeting therapeutic substances on natural fibrin, for use
PT in affinity purificn. of polypeptide(s) etc.

XX Claim 11; Fig 3A; 32pp; English.

XX The product may be expressed from a transformed micro-organism, esp.

CC E. coli. FN binds to fibrin, heparin and Staphylococcus aureus, and may be
CC used to target a therapeutic agent onto natural fibrin eg. a blood clot.
CC It may also be used in affinity purification of a polypeptide, conjugated
CC to the collagen binding site of FN and immobilised on a collagen surface.
CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 2327 AA;

Query Match 100.0%; Score 600; DB 1; Length 2327;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQMVQPSQSPVAVSQSGKPCYDNGKHQIINQWERTYLGVLVCTCYGSGRGFNCESKP 60
Db 4 QAAQMVQPSQSPVAVSQSGKPCYDNGKHQIINQWERTYLGVLVCTCYGSGRGFNCESKP 63

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106

Db 64 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 109

RESULT 7

AAR15468
ID AAR15468 standard; protein; 2327 AA.

XX AAR15468;

XX 25-MAR-2003 (revised)

DT 12-MAR-1992 (first entry)

XX Human fibronectin.

XX Fibrin-imaging; atherosclerosis; thrombus inhibitor.

XX Homo sapiens.

XX WO9117765-A.

XX 28-NOV-1991.

XX 21-MAY-1990; 90US-00526397.

XX 21-MAY-1990; 90US-00526397.

XX (BIOT-) BIO-TECHN GEN CORP.

XX Vogel T, Levanon A, Werber M, Guy R, Panet A, Hartman J;

XX Shaked H;

XX WPI; 1991-369004/50.

XX N-PSDB; AAQ15214.

XX New fibrin binding domain polypeptide(s) - useful in imaging fibrin-
PT contg. substances, to inhibit thrombus formation and treat wounds.

XX Disclosure; Fig 2; 191pp; English.

XX The amino acid sequence is that of human fibronectin, this can be used to
CC derive polypeptides which are identical to part of the fibrin- binding
CC domain (FBD) of fibronectin. These polypeptides can be used to inhibit
CC thrombus formation; or (coupled to a thrombolytic agent) to induce
CC thrombolysis, or to treat wounds, e.g. in skin, eyes or tendons (in
CC conjunction with a polypeptide which includes a part of the cell-binding
CC domain (CBD) of fibronectin). These polypeptides are easier to prepare
CC than the full 31kD polypeptide. It can also be used to image fibrin-
CC contg. materials, esp. a thrombus or atherosclerotic plaque, pref.
CC using a gamma counter. (Updated on 25-MAR-2003 to correct FI field.)

XX SQ Sequence 2327 AA;

Query Match 100.0%; Score 600; DB 2; Length 2327;

Best Local Similarity 100.0%; Pred. No. 2.5e-52;

Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 100.0%; Score 600; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCSEK 60
DB 5 QAOQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCSEK 64

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 110

RESULT 10
ABR41106
ID ABR41106 standard; protein; 2328 AA.
XX ABR41106;
AC
XX
XX
XX 02-JUN-2003 (first entry)
XX Human fibronectin gene FN protein product.
DE
XX
XX Human; high bone mass; HBM; LRP5; LRP6; transgenic; bone mass modulation;
KW gene therapy; bone density modulation; bone strength; trabecular number;
KW bone size; bone tissue connectivity; bone disease; osteoporosis;
KW osteomalacia; rickets; Paget's disease; neoplasm of the bone.
XX
XX Homo sapiens.
OS
XX
XX WO200292764-A2.
FN
XX
XX 21-NOV-2002.
PD
XX
XX 13-MAY-2002; 2002WO-US014876.
PF
XX
XX 11-MAY-2001; 2001US-0290071P.
PR
XX 17-MAY-2001; 2001US-0291311P.
PR
XX 01-FEB-2002; 2002US-0353058P.
PR
XX 04-MAR-2002; 2002US-0361293P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX (AMHP) WYETH.
PA
XX
XX Babi J P, Bex FJ, Yaworsky PJ, Bodine PV;
FI
XX
XX WPI; 2003-129278/12.
DR
XX
XX New transgenic animals (e.g. mice), useful as models for studying bone
PT density modulation, developing drugs for treating or preventing bone
PT diseases (e.g. osteoporosis), or diagnosing diseases characterized by
PT reduced bone density.
XX
XX Disclosure; Page 532-538; 603pp; English.

CC The invention relates to novel transgenic animals expressing the high
CC bone mass (HBM) gene, expressing the corresponding wild type HBM gene,
CC comprising an alteration of the gene encoding LRP5 or LRP6, or expressing
CC an LRP5 that is modulated by an altered gene control sequence introduced
CC by homologous or non-homologous recombination. The transgenic animals are
CC for the study of bone density modulation or bone mass modulation. The
CC invention has osteopathic and cytostatic activity. The polynucleotides of
CC the invention may have a use in gene therapy. The transgenic animals and
CC nucleic acids are for the study of bone density modulation, where the
CC bone mass is modulated relative to non-transgenic animals of the same
CC species in more than one parameter selected from bone density, bone
CC strength, trabecular number, bone size, or bone tissue connectivity. The
CC transgenic animals, nucleic acids and methods are useful for identifying
CC molecules involved in bone development, and for developing pharmaceutical
CC compositions, which may be employed for treating or preventing bone
CC diseases, e.g. osteoporosis, osteomalacia, rickets, Paget's disease, or
CC neoplasms of the bone. The transgenic animals and nucleic acids are also
CC useful in methods for diagnosing diseases involved in bone development,

CC or characterised by reduced bone density or mass. The present sequence is
CC used in the exemplification of the invention

XX
XX
XX Sequence 2328 AA;

Query Match 100.0%; Score 600; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCSEK 60
DB 5 QAOQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNFNCSEK 64

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 110

RESULT 11
ABR92078
ID ABR92078 standard; protein; 2328 AA.
XX ABR92078;
AC
XX
XX 10-SEP-2003 (first entry)
XX Human cervical cancer cell marker protein SEQ ID NO:64.
DE
XX
XX Human; cervical cancer; cervical cancer marker; cancer therapy;
KW detection; gene therapy; vaccine.
KW
XX
XX Homo sapiens.
OS
XX
XX WO2002101075-A2.
FN
XX
XX 19-DEC-2002.
PD
XX
XX 12-JUN-2002; 2002WO-US018638.
PF
XX
XX 13-JUN-2001; 2001US-0298155P.
PR
XX 13-JUN-2001; 2001US-0298159P.
PR
XX 14-NOV-2001; 2001US-0335936P.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
FI Gannavarapu M, Glatt K, Hoersch S;
PI
XX
XX WPI; 2003-156967/15.
DR
XX N-PSDB; ACF12859.
DR
XX
XX New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
PT
XX
XX Claim 4; Page 212-217; 386pp; English.

CC ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials

```
SQ Sequence 2328 AA;
Query Match 100.0%; Score 600; DB 6; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db 5 QAOQWVQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 110

RESULT 12
ADB70378
ID ADB70378 standard; protein; 2328 AA.
XX AC
XX ADB70378;
XX DT 04-DEC-2003 (first entry)
XX DE Fibronectin SEQ ID NO:70.
XX KW cancer; malignant pleural mesothelioma; MPW; lung adenocarcinoma;
XX KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
XX KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
XX KW human.
XX OS Homo sapiens.
XX PN WO2003021229-A2.
XX PD 13-MAR-2003.
XX PF 05-SEP-2002; 2002WO-US028203.
XX PR 05-SEP-2001; 2001US-0317389P.
XX PR 30-AUG-2002; 2002US-00236031.
XX PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX PI Gordon GJ, Jensen RV, Gullans SR, Bueno R;
XX WPI; 2003-290233/28.
XX DR N-PSDB; ADB70377.
XX PT Diagnosing cancer cells in tissue sample, or determining prognosis or
XX PT outcome of cancer patient, by calculating ratio of expression levels of
XX PT genes that are differentially expressed in cancer and non cancer tissues.
XX PS Claim 77; Page 369-376; 396pp; English.
XX CC The present invention describes a method (M1) for diagnosing the presence
XX CC of cancer cells or non-cancer cells in a tissue sample, or determining
XX CC the prognosis or outcome of a cancer patient. M1 involves providing a set
XX CC of genes that are differentially expressed in cancerous or non-cancerous
XX CC conditions, determining the expression levels of the set of genes and
XX CC calculating a ratio of the expression levels of the differentially
XX CC expressed genes. M1 is useful for diagnosing the presence of cancer cells
XX CC or non-cancer cells in a tissue sample, where the cancer is malignant
XX CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma,
XX CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell
XX CC lymphoma, follicular lymphoma and ovarian cancer, and for determining
XX CC prognosis or outcome of a cancer patient. The ratio of expression levels
XX CC of differentially expressed genes is used as an indicator of cancer type,
XX CC cancer class, and/or cancer prognosis, all of which are useful for
XX CC determining a course of treatment of a patient. The present sequence
XX CC represents a human protein which is used in an example from the present
XX CC invention.
XX SQ Sequence 2328 AA;

Query Match 100.0%; Score 600; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db 5 QAOQWVQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 64

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
Db 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 110

RESULT 13
ADB98726
ID ADB98726 standard; protein; 2328 AA.
XX AC
XX ADB98726;
XX DT 04-DEC-2003 (first entry)
XX DE Human fibronectin.
XX KW Osteopathic; Gene therapy; High Bone Mass; HBM; LRP5; Zmaxi; LRP6;
XX KW bone mass modulation; osteoporosis.
XX OS Homo sapiens.
XX PN WO200292000-A2.
XX PD 21-NOV-2002.
XX PF 13-MAY-2002; 2002WO-US014877.
XX PR 11-MAY-2001; 2001US-0290071P.
XX PR 17-MAY-2001; 2001US-0291311P.
XX PR 01-FEB-2002; 2002US-0353058P.
XX PR 04-MAR-2002; 2002US-0361293P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PA (AMGP) WYETH.
XX PI Allen K, Anisowicz A, Graham JR, Morales A, Yaworsky PJ, Liu W;
XX WPI; 2003-129214/12.
XX DR P-PSDB; ADB98703.
XX PT New nucleic acid comprising a mutation in LRP5 or LRP6, useful for
XX PT diagnosing a HBM-like phenotype in a subject and for preparing a
XX PT composition for modulating bone mass and/or lipid levels in a subject
XX PT suffering from e.g. osteoporosis.
XX PS Disclosure; SEQ ID NO 677; 629pp; English.
XX CC The present invention relates to High Bone Mass (HBM), LRP5 (Zmaxi) and
XX CC LRP6 mutants, which results in a HBM-like phenotype when expressed in a
XX CC cell. The HBM-like phenotype results in bone mass modulation and/or lipid
XX CC level modulation. The invention is useful for diagnosing a HBM-like
XX CC phenotype in a subject and for preparing a composition for modulating
XX CC bone mass and/or lipid levels in a subject suffering from e.g.
XX CC osteoporosis. The present sequence was used to illustrate the invention.
XX SQ Sequence 2328 AA;

Query Match 100.0%; Score 600; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQWVQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 60
Db 5 QAOQWVQSPVAVSQSKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGNFNCESKP 64
```

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
ADJ37157
DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 110

RESULT 14

ADJ37157
ID ADE82522
XX ADE82522 standard; protein; 2328 AA.

AC ADE82522;

XX 29-JAN-2004 (first entry)

XX Human protein sequence related to the invention #12.

KW LRP5; LRP6; HBM; Dkk activity; Osteopathic; Antiinflammatory;
KW Antiarthritic; bone mass disorders; osteoporosis; hypercalcaemia;
KW hyperostosis; osteogenesis; Wnt signaling.

XX Homo sapiens.

XX WO200292015-A2.

XX 21-NOV-2002.

XX 17-MAY-2002; 2002WO-US015982.

XX 17-MAY-2001; 2001US-0291311P.

XX 01-FEB-2002; 2002US-0353058P.

XX 04-MAR-2002; 2002US-0361293P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX (AWHP) WYETH.

XX Allen K, Anisowicz A, Bhat BM, Damagnez V, Robinson JA;

XX Yaworsky PJ;

XX WPI; 2003-129219/12.

XX Regulating LRP5, LRP6 or HBM activity in a subject, useful for modulating
PT lipid levels and/or bone mass, and for in treating bone mass disorders,
PT e.g. osteoporosis, comprises administering a composition which modulates
PT a Dkk activity.

XX Disclosure; SEQ ID NO 98; 173pp; English.

XX The present invention relates to regulating LRP5, LRP6 or HBM activity in
CC a subject comprising administering a composition which modulates a Dkk
CC activity. The method is useful for modulating lipid levels and/or bone
CC mass, and is useful in treating or diagnosing abnormal lipid levels and
CC bone mass disorders, such as osteoporosis, bone fracture, age-related
CC loss of bone, a chondrodystrophy, drug-induced bone disorder, high bone
CC turnover, hypercalcaemia, hyperostosis, osteogenesis, imperfecta,
CC osteomalacia, osteomyelitis, Paget's disease, osteoarthritis, and
CC rickets. Modulators of Dkk activity are useful for as reagents in
CC studying bone mass and lipid level modulation, in modulating wnt
CC signaling, or treating Dkk-mediated disorders. The present sequence
CC represents a human protein sequence related to the invention.

XX Sequence 2328 AA;

Query Match 100.0%; Score 600; DB 7; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQQMVQPSVAVSQSPGCGYDNGKHQYIQINQWERTYLGNIWLCTCYGSGRGFNCSKP 60

DB 5 QAQQMVQPSVAVSQSPGCGYDNGKHQYIQINQWERTYLGNIWLCTCYGSGRGFNCSKP 64

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 110

RESULT 15

ADJ37157
ID ADJ37157 standard; protein; 2328 AA.

XX AC ADJ37157;

XX 22-APR-2004 (first entry)

XX Human malignant pleural mesothelioma (MPM) protein #31.

XX Human; malignant pleural mesothelioma; MPM; tumour; lung adenocarcinoma;
KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
KW cytostatic.

XX Homo sapiens.

XX US2003219760-A1.

XX 27-NOV-2003.

XX 05-SEP-2002; 2002US-00236031.

XX 05-SEP-2001; 2001US-0317389P.

XX 30-AUG-2002; 2002US-0407431P.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Gordon GJ, Jensen RV, Gullans SR, Bueno R;

XX WPI; 2004-141744/14.

XX N-PSDB; ADJ37156.

XX Diagnosing the presence of cancer or non-cancer cells in tissue sample,
PT useful for diagnosing malignant pleural mesothelioma comprises
PT determining ratio of expression level of a set of genes expressed in
PT cancer tissues.

XX Claim 77; SEQ ID NO 70; 53pp; English.

XX The invention relates to a method of diagnosing the presence of cancer
CC cells or non-cancer cells in a tissue sample, determining prognosis or
CC outcome of a cancer patient, selecting a course of treatment for a
CC subject having or suspected of having malignant pleural mesothelioma
CC (MPM) and evaluating treatment of MPM comprising determining the ratio of
CC the expression level of a set of genes differentially expressed in a
CC cancer tissue. The cancer is chosen from MPM, lung adenocarcinoma,
CC squamous carcinoma, medulloblastoma, prostate cancer, breast cancer,
CC diffuse large B-cell lymphoma, follicular lymphoma and ovarian cancer.
CC The method is useful for diagnosing MPM in a subject suspected of having
CC MPM which involves obtaining a tissue sample suspected of being cancerous
CC from a subject and determining the expression of nucleic acid markers or
CC its expression products in the tissue sample. This sequence represents a
CC human MPM protein of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 2328 AA;

Query Match 100.0%; Score 600; DB 8; Length 2328;
Best Local Similarity 100.0%; Pred. No. 2.5e-52;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQQMVQPSVAVSQSPGCGYDNGKHQYIQINQWERTYLGNIWLCTCYGSGRGFNCSKP 60

DB 5 QAQQMVQPSVAVSQSPGCGYDNGKHQYIQINQWERTYLGNIWLCTCYGSGRGFNCSKP 64

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

DB 65 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 110

Search completed: November 3, 2004, 23:49:25
Job time : 54.5479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:42:10 ; Search time 10.8904 Seconds
(without alignments)
936.510 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQQMVQPSPVAVSQSPKG.....SMIWDCTCAGRGGRISCTI 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	100.0	2386	1 FNHU	fibronectin precu
2	575	95.8	2267	1 FNBO	fibronectin - bovi
3	574	95.7	2477	2 S14428	fibronectin precu
4	492	82.0	2481	2 A43908	fibronectin - Afri
5	164	27.3	103	2 A43173	fibronectin - Afri
6	141	23.5	1020	2 A29355	fibronectin - chic
7	138	23.0	190	2 I51279	fibronectin - east
8	80	13.3	732	2 I52361	testicular metallo
9	79.5	13.2	2524	2 A35844	Xotch protein - Af
10	77.5	12.9	477	2 JS0597	t-plasminogen acti
11	77	12.8	473	2 A56175	adhesive plaque pr
12	76	12.7	328	4 A58437	hypothetical cyste
13	76	12.7	713	2 I65253	disintegrin-like t
14	76	12.7	5147	1 IJFRTM	cadherin-related t
15	74.5	12.4	5376	2 T42215	zonadhesin - mouse
16	74	12.3	1964	2 T09059	notch4 - mouse
17	73.5	12.2	559	1 A29941	t-plasminogen acti
18	73	12.2	655	1 A46688	hepatocyte growth
19	72.5	12.1	2871	2 A55624	fibrillin-1 precu
20	72	12.0	955	2 A45441	thrombospondin 4 -
21	72	12.0	2139	2 A35672	crumbs protein - f
22	72	12.0	2437	2 S42612	transmembrane prot
23	71.5	11.9	1064	2 A40136	fibropellin Ia - s
24	71.5	11.9	2555	2 A40043	notch protein homo
25	71	11.8	257	2 T33292	hypothetical prote
26	70.5	11.8	838	2 T20130	hypothetical prote
27	70.5	11.8	832	2 T20125	hypothetical prote
28	70.5	11.8	957	2 T15976	hypothetical prote
29	70.5	11.8	1297	2 T30274	proteolisin - se

ALIGNMENTS

RESULT 1

FNHU

fibronectin precursor [validated] - human

N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A92

R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656

A;Accession: A26460

A;Molecule type: DNA

A;Residues: 1-49 <DEA>

A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:g182686; PIDN:AAA533

R;Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A;Title: Evolution of the fibronectin gene.

A;Reference number: A26284; MUID:86111901; PMID:3003095

A;Accession: A26284

A;Molecule type: DNA

A;Residues: 1447-1540 <OLD>

A;Cross-references: GB:M12549; NID:g182688

A;Note: the authors translated the codon TTC for residue 1494 as Glu

R;Paolella, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A;Reference number: S00848; MUID:88233940; PMID:3375063

A;Accession: S03917

A;Molecule type: DNA

A;Residues: 1594-1767, 'V', 1769-1783 <PAO>

A;Cross-references: EMBL:X07718; NID:g31402

A;Note: the authors translated the codon AAC for residue 1631 as Asp

R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene

A;Reference number: A24854; MUID:87030929; PMID:3770201

A;Accession: A24854

A;Molecule type: DNA

A;Residues: 1992-2147 <VIB>

A;Cross-references: GB:X04530; NID:g31436

R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.

FEBS Lett. 207, 145-148, 1986

A;Title: Human fibronectin is synthesized as a pre-propolypeptide.

A;Reference number: A24476; MUID:87030890; PMID:3770189

A;Accession: A24476

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-14, 'Q', 16-38 <GUT>

R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1755-1759, 1985

A;Title: Primary structure of human fibronectin: differential splicing may generate at 1

A;Reference number: A91008; MUID:85284965; PMID:2392939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080,2112-2386 <KOR>
A;Cross-references: GB:X02761
R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptides
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 973-2080,2112-2386 <K02>
A;Cross-references: GB:X00739
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:K00055; NID:gl82680; PIDN:AA52459.1; PID:gl82683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with rat
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386 <BER>
A;Cross-references: GB:M10905; NID:gl82696; PIDN:AA52462.1; PID:gl82697
R;Umezawa, K.; Kornblihtt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UMB>
A;Cross-references: GB:M27589; NID:gl82705; PIDN:AA52465.1; PID:gl82706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991,2017-2039 <UM2>
A;Cross-references: GB:M27590
R;Sekiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: 152394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SEK>
A;Cross-references: GB:M14060; NID:gl82701; PIDN:AA52464.1; PID:gl82704
R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <K03>
A;Cross-references: GB:K00799; NID:gl82681; PIDN:AA52460.1; PID:gl82684
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the p
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300,551-560 <GAR2>

R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GRI>
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: Residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87241275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
R;Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080,2112-2356 <GAR4>
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,
ation, and transformation.
C;Genetics:
A;Gene: GDB:FN1
A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-2q34
A;Intons: 49/3; 1286/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;97-135/Domain: fibronectin type I repeat homology <1F1>
F;141-179/Domain: fibronectin type I repeat homology <1F2>
F;186-225/Domain: fibronectin type I repeat homology <1F3>
F;231-270/Domain: fibronectin type I repeat homology <1F4>
F;308-608/Domain: fibronectin type I repeat homology <1F5>
F;308-342/Domain: collagen binding <CBR>
F;360-401/Domain: fibronectin type II repeat homology <2F1>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <1F7>
F;518-555/Domain: fibronectin type I repeat homology <1F8>
F;561-599/Domain: fibronectin type I repeat homology <1F9>
F;609-692/Domain: fibronectin type III repeat homology <3FA>
F;616-706/Domain: heparin binding <HPB>
F;719-801/Domain: fibronectin type III repeat homology <3FB>
F;810-891/Domain: fibronectin type III repeat homology <3FC>

F;906-988/Domain: fibronectin type III repeat homology <3FD>
 F;996-1077/Domain: fibronectin type III repeat homology <3FB>
 F;1086-1164/Domain: fibronectin type III repeat homology <3FF>
 F;1173-1258/Domain: fibronectin type III repeat homology <3FG>
 F;1266-1349/Domain: fibronectin type III repeat homology <3FH>

Query Match 100.0%; Score 600; DB 1; Length 2386;
 Best Local Similarity 100.0%; Pred. No. 1.7e-51;
 Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSVAVSQSGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
 |||||
 Db 32 QAOQMVQPSVAVSQSGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 91
 |||||

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCITCIGAGRGRIISCTI 106
 |||||
 Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCITCIGAGRGRIISCTI 137
 |||||

RESULT 2
 FNBO
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
 C;Accession: A26452; B21165; A23292
 R;Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.
 Eur. J. Biochem. 161, 441-453, 1986
 A;Title: Complete primary structure of bovine plasma fibronectin.
 A;Reference number: A26452; MUID:87054047; PMID:3780752
 A;Accession: A26452
 A;Molecule type: protein
 A;Residues: 1-2265 <SKO>
 A;Cross-references: UNIPROT:P07589
 R;Kornblihtt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
 Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
 A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
 A;Reference number: A21165; MUID:83221567; PMID:6304699
 A;Accession: B21165
 A;Molecule type: mRNA
 A;Residues: 2170-2265 <KOR>
 A;Cross-references: GB:K00800; NID:gl63055; PIDN:AAA30521.2; PID:g5713323
 R;Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sotth
 Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
 A;Title: Partial primary structure of bovine plasma fibronectin: three types of internal
 A;Reference number: A23292; MUID:83117805; PMID:6218503
 A;Accession: A23292
 A;Molecule type: protein
 A;Residues: 1-16,'C',18-20,'S',22-432;447-463;1367-1517;1567-1673;2062-2176,'N',2178-226
 C;Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.
 C;Comment: The plasma fibronectin molecule consists of two chains, which are connected b
 C;Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
 aling, and maintenance of cell shape.
 C;Comment: Plasma fibronectin is synthesized by hepatocytes.
 C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
 C;Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu
 F;21-241/Domain: fibrin and heparin binding <FBR>
 F;21-56/Domain: fibronectin type I repeat homology <1F1>
 F;66-104/Domain: fibronectin type I repeat homology <1F2>
 F;110-148/Domain: fibronectin type I repeat homology <1F3>
 F;155-194/Domain: fibronectin type I repeat homology <1F4>
 F;200-239/Domain: fibronectin type I repeat homology <1F5>
 F;277-577/Domain: collagen binding <CBR>
 F;277-311/Domain: fibronectin type I repeat homology <1F6>
 F;329-370/Domain: fibronectin type II repeat homology <2F1>
 F;389-430/Domain: fibronectin type II repeat homology <2F2>
 F;439-477/Domain: fibronectin type I repeat homology <1F7>
 F;487-524/Domain: fibronectin type I repeat homology <1F8>
 F;530-568/Domain: fibronectin type I repeat homology <1F9>
 F;578-661/Domain: fibronectin type III repeat homology <FN3A>
 F;688-770/Domain: fibronectin type III repeat homology <FN3B>
 F;779-860/Domain: fibronectin type III repeat homology <FN3C>
 F;875-957/Domain: fibronectin type III repeat homology <FN3D>
 F;965-1046/Domain: fibronectin type III repeat homology <FN3E>

F;1055-1134/Domain: fibronectin type III repeat homology <FN3F>
 F;1142-1227/Domain: fibronectin type III repeat homology <FN3G>
 F;1235-1318/Domain: fibronectin type III repeat homology <FN3H>
 F;1326-1404/Domain: fibronectin type III repeat homology <GN3I>
 F;1410-1517/Domain: cell attachment <CAD>
 F;1416-1502/Domain: fibronectin type III repeat homology <FN3J>
 F;1493-1495/Region: cell attachment (R-G-D) motif
 F;1510-1592/Domain: fibronectin type III repeat homology <FN3K>
 F;1600-1870/Domain: heparin binding <HB2>
 F;1600-1682/Domain: fibronectin type III repeat homology <FN3L>
 F;1692-1773/Domain: fibronectin type III repeat homology <FN3M>
 F;1781-1863/Domain: fibronectin type III repeat homology <FN3N>
 F;1970-1972/Region: cell attachment (R-G-D) motif
 F;1982-2062/Domain: fibronectin type III repeat homology <FN3O>
 F;1985-2216/Domain: fibrin binding <FB2>
 F;2085-2124/Domain: fibronectin type I repeat homology <1F10>
 F;2130-2167/Domain: fibronectin type I repeat homology <1F11>
 F;2174-2209/Domain: fibronectin type I repeat homology <1F12>
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F;3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experime
 F;21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-
 7,2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted
 F;399,497,511,846,976,1213,1987/Binding site: carbohydrate (Asn) (covalent) #status exp
 F;1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent
 F;1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F;2246/Disulfide bonds: interchain (to 2250) #status predicted
 F;2250/Disulfide bonds: interchain (to 2246) #status predicted
 F;2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 95.8%; Score 575; DB 1; Length 2265;
 Best Local Similarity 94.3%; Pred. No. 5e-49;
 Matches 100; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QAOQMVQPSVAVSQSGKCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
 |||||
 Db 1 QAOQMVQPSVAVSQSGKCYDNGKHQYQINQOWERTYLGNSALVCTCYGSGRGFNCSKP 60
 |||||

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCITCIGAGRGRIISCTI 106
 |||||
 Db 61 EPEETCFDKYTGNTYRVGDTYERPKDSMIWDCITCIGAGRGRIISCTI 106
 |||||

RESULT 3
 S14428
 fibronectin precursor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
 C;Accession: S14428; S12455; A22319; S46203; S00459; A27352; I59049
 R;Hynes, R.O.
 submitted to the EMBL Data Library, July 1989
 A;Reference number: S14428
 A;Accession: S14428
 A;Molecule type: mRNA
 A;Residues: 1-2477 <HYN>
 A;Cross-references: UNIPROT:P04937; EMBL:X15906; NID:G56163; PIDN:CAA34020.1; PID:g56163
 R;Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
 EMBO J. 6, 2573-2580, 1987
 A;Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
 A;Reference number: S12455; MUID:88054951; PMID:2445560
 A;Accession: S12455
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 609-1810,'T',1812-2283 <SCH>
 A;Cross-references: EMBL:X15906
 R;Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
 Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
 A;Title: A single rat fibronectin gene generates three different mRNAs by alternative sp
 A;Reference number: A22319; MUID:84238037; PMID:6089177
 A;Accession: A22319
 A;Molecule type: DNA
 A;Residues: 2052-2237 <TAM>
 R;Falkenberg, C.; Enghild, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstroem, B.
 Biochem. J. 301, 745-751, 1994

uctive monkey tissues.

C:Reference number: I52361; MUID:96077150; PMID:7492319

A:Accession: I52361

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-732 <RES>

C:Cross-references: UNIPROT:Q28484; EMBL:X87205; NID:G1061158; PIDN:CAA60663.1; PID:G106

C:Genetics:

A:Gene: tMDC IVa

C:Superfamily: mouse meltrin alpha; disintegrin homology

F:403-485/Domain: disintegrin homology <DIS>

Query Match 13.3%; Score 80; DB 2; Length 732;

Best Local Similarity 26.5%; Pred. No. 3.4;

Matches 35; Conservative 10; Mismatches 45; Indels 42; Gaps 5;

QY 2 AQQVQSPVAVSOSKPGCVNDGKHQYQINQWERTYLGVLVCTCYGSGRG-----FN 55

DB 97 ALQVDDPVP-----PDCYVLG-----YLEEVPLSMVTVDTCYGLRGIMKLLDLIA 142

QY 56 CESKPEAEETCFD-----KYTGNTYRVGDITYE-----RPKDSMIWDCT 93

DB 143 YEIKPLQDSRFEHVVSQIVAEPNATGPTFRGDNEETDPLFSEANDMNPRISTLLYSS 202

QY 94 CIGAGRGRICT 105

DB 203 HRGNIKHVQCS 214

RESULT 9

A35844

C:Species: Xenopus laevis (African clawed frog)

C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 16-Aug-2004

C:Accession: A35844

R:Coffman, C.; Harris, W.; Kintner, C.

A:Title: Xotch, the Xenopus homolog of Drosophila notch.

A:Reference number: A35844; MUID:90385285; PMID:2402639

A:Accession: A35844

A:Status: preliminary; mRNA

A:Molecule type: mRNA

A:Residues: 1-2524 <COF>

C:Superfamily: Notch protein; ankyrin repeat homology; EGF homology

F:146-177/Domain: EGF homology <EGX1>

F:184-215/Domain: EGF homology <EGF1>

F:222-254/Domain: EGF homology <EGF>

F:456-487/Domain: EGF homology <EGX2>

F:757-788/Domain: EGF homology <EGF3>

F:1025-1056/Domain: EGF homology <EGX3>

F:1924-1956/Domain: ankyrin repeat homology <AN1>

F:1957-1989/Domain: ankyrin repeat homology <AN2>

F:1991-2023/Domain: ankyrin repeat homology <AN3>

F:2024-2056/Domain: ankyrin repeat homology <AN4>

F:2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match 13.2%; Score 79.5; DB 2; Length 2524;

Best Local Similarity 26.8%; Pred. No. 12;

Matches 26; Conservative 10; Mismatches 26; Indels 35; Gaps 5;

QY 17 SKPCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNCESKPE-----AETCTPDK 69

DB 609 SKP-CLANG-----QCTDRENGYICTCPKGTGVNCTKIDDCASNLCDNGKCIDK 658

QY 70 YTGNTYRVGDFYRPKDSMIWDCTCIGAGRGRICTI 106

DB 659 IDG-----YECTCEPYTGKL-CNI 677

RESULT 10

JS0597

t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat

N;Alternate names: tissue plasminogen activator

C:Species: Desmodus rotundus (common vampire bat)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C:Accession: JS0597

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-732 <RES>

C:Cross-references: UNIPROT:Q28484; EMBL:X87205; NID:G1061158; PIDN:CAA60663.1; PID:G106

C:Genetics:

A:Gene: tMDC IVa

C:Superfamily: mouse meltrin alpha; disintegrin homology

F:403-485/Domain: disintegrin homology <DIS>

Query Match 13.3%; Score 80; DB 2; Length 732;

Best Local Similarity 26.5%; Pred. No. 3.4;

Matches 35; Conservative 10; Mismatches 45; Indels 42; Gaps 5;

QY 2 AQQVQSPVAVSOSKPGCVNDGKHQYQINQWERTYLGVLVCTCYGSGRG-----FN 55

DB 97 ALQVDDPVP-----PDCYVLG-----YLEEVPLSMVTVDTCYGLRGIMKLLDLIA 142

QY 56 CESKPEAEETCFD-----KYTGNTYRVGDITYE-----RPKDSMIWDCT 93

DB 143 YEIKPLQDSRFEHVVSQIVAEPNATGPTFRGDNEETDPLFSEANDMNPRISTLLYSS 202

QY 94 CIGAGRGRICT 105

DB 203 HRGNIKHVQCS 214

Query Match 12.9%; Score 77.5; DB 2; Length 477;

Best Local Similarity 29.3%; Pred. No. 4;

Matches 24; Conservative 12; Mismatches 35; Indels 11; Gaps 5;

QY 15 SOSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNCESKPEAETCFDKYTGNT 74

DB 86 SCSEPRCFNGG-----TCWQAVYFSD-FVQCQPGAGYTKRCCEVDIFA--TCYEG-QGVT 135

QY 75 YR-VGDTYRPRKDSMIWDCTCI 95

DB 136 YRGTWSTAESRVEECINWNSLL 157

RESULT 11

A56175

C:Species: Mytilus galloprovincialis (Mediterranean mussel)

C:Date: 27-Apr-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004

C:Accession: A56175

R:Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.

J. Biol. Chem. 270, 6698-6701, 1995

A:Title: Mussel adhesive plaque protein gene is a novel member of epidermal growth factor

A:Reference number: A56175; MUID:95204464; PMID:7896812

A:Accession: A56175

A:Molecule type: mRNA

A:Residues: 1-473 <INO>

C:Cross-references: UNIPROT:Q25464; GB:D43794; NID:G602767; PIDN:BAA07852.1; PID:d1008433

C:Keywords: duplication

F:1-17/Domain: signal sequence #status predicted <SIG>

F:387-419/Domain: EGF homology <EGF1>

F:429-460/Domain: EGF homology <EGF>

F:23,36,43,56,75,382,424,455,468,473/Modified site: 3',4'-dihydroxyphenylalanine (Tyr)

Query Match 12.8%; Score 77; DB 2; Length 473;

Best Local Similarity 30.3%; Pred. No. 4.4;

Matches 27; Conservative 5; Mismatches 39; Indels 18; Gaps 4;

QY 18 KPG-CYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNCESKPEAETCFDKYTGNTYR 76

DB 307 KPNPCQNRGRCYPDNSD-----DGFKRCVGVGKGTCEKPNPCNTKPCNKGKCN 359

QY 77 VGDYRPRKDSMIWDCTCIGAGRGRICT 105

DB 360 NG-----KIYTCCKAYGWRGR-HCT 378

RESULT 12

A58437
hypotheoretical cysteine-rich protein IVC, processed pseudogene - crab-eating macaque
N;Alternate names: disintegrin-like testicular metalloproteinase IVC
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 13-Dec-1996 #sequence revision 31-Dec-1996 #text_change 09-Jul-2004
C;Accession: A58437; S59854; I65254
R;Perry, A.C.F.; Jones, R.; Hall, L.
Biochem. J. 312, 239-244, 1995
A;Title: Analysis of transcripts encoding novel members of the mammalian metalloproteases active monkey tissues.
A;Reference number: I52361; MUID:96077150; PMID:7492319
A;Accession: A58437
A;Status: translated from GB/EMBL/DBJ; conceptual translation of pseudogene
A;Molecule type: mRNA
A;Residues: 1-328 <PER>
A;Cross-references: UNIPROT:Q28486; EMBL:X87207; NID:g1061162; PIDN:CRA60665.1; PID:g1061162
A;Experimental source: adult testis cDNA library
A;Note: submitted to the EMBL Data Library, May 1995
C;Genetics:
A;Gene: tMDC IVC
C;Keywords: pseudogene

Query Match .. 12.7%; Score 76; DB 4; Length 328;
Best Local Similarity 29.4%; Pred. No. 3.9; Mismatches 5; Indels 36; Gaps 5;
Matches 32; Conservative 5

Qy 2 AQQVQPSQSPVAVSQSKPGCYNGKHQYINQWERTYLGVLVCTCYGSGRG-----FN 55
Db 97 ALQVDDPYIP-----PCYYLG-----YLEEVLPSMVTVDTCYGGLRGIMKLDLLA 142

Qy 56 CSKPEAETCFD-----KYTGNTYRVGDT-----YRPRKDSM 88
Db 143 YEIKPLQDSRRRHHVVSQIVAEPNATGPTFRDGDNEETDPLFSEANDSM 191

RESULT 13
I65253
disintegrin-like testicular metalloproteinase (EC 3.4.24.-) IVb - crab-eating macaque (H
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I65253
R;Perry, A.C.F.; Jones, R.; Hall, L.
Biochem. J. 312, 239-244, 1995
A;Title: Analysis of transcripts encoding novel members of the mammalian metalloproteases active monkey tissues
A;Reference number: I52361; MUID:96077150; PMID:7492319
A;Accession: I65253
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-713 <RES>
A;Cross-references: UNIPROT:Q28485; EMBL:X87206; NID:g1061160; PIDN:CRA60664.1; PID:g1061160
C;Genetics:
A;Gene: tMDC IVb
C;Superfamily: mouse meltrin alpha; disintegrin homology
C;Keywords: hydrolase; metalloproteinase
F;384-466/Domain: disintegrin homology <DIS>

Query Match 12.7%; Score 76; DB 2; Length 713;
Best Local Similarity 29.4%; Pred. No. 8.2;
Matches 32; Conservative 5; Mismatches 36; Indels 36; Gaps 5;
Matches 32; Conservative 5

Qy 2 AQQVQPSQSPVAVSQSKPGCYNGKHQYINQWERTYLGVLVCTCYGSGRG-----FN 55
Db 78 ALQVDDPYIP-----PCYYLG-----YLEEVLPSMVTVDTCYGGLRGIMKLDLLA 123

Qy 56 CSKPEAETCFD-----KYTGNTYRVGDT-----YRPRKDSM 88
Db 124 YEIKPLQDSRRRHHVVSQIVAEPNATGPTFRDGDNEETDPLFSEANDSM 172

RESULT 14
IUFFTM
cadherin-related tumor suppressor precursor - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Feb-1997
C;Accession: A41087; B41087
R;Mahoney, P.A.; Weber, U.; Onofrechuk, P.; Biessmann, H.; Bryant, P.J.; Goodman, C.S.
Cell 67, 853-868, 1991
A;Title: The fat tumor suppressor gene in Drosophila encodes a novel member of the cadherin superfamily.
A;Reference number: A41087; MUID:92069752; PMID:1959133
A;Accession: A41087
A;Molecule type: mRNA
A;Residues: 143-485;1279-5147 <MAH>
A;Cross-references: GB:M80537
A;Accession: B41087
A;Molecule type: DNA
A;Residues: 1-142;487-1278 <MA2>
A;Cross-references: GB:M80537
A;Note: 1229-Gly and 1233-Ser were also found
C;Genetics:
A;Gene: fat
A;Cross-references: FlyBase:FBgn0001075
C;Superfamily: cadherin-related tumor suppressor; cadherin repeat homology; EGF homology
C;Keywords: calcium binding; cell adhesion; duplication; transmembrane protein
F;1-35/Domain: signal sequence #status predicted <SIG>
F;36-5147/Product: cadherin-related tumor suppressor #status predicted <MAT>
F;36-483/Domain: extracellular #status predicted <EXT>
F;51-156/Domain: cadherin repeat homology <CR1>
F;159-270/Domain: cadherin repeat homology <CR2>
F;271-382/Domain: cadherin repeat homology <CR3>
F;390-494/Domain: cadherin repeat homology <CR4>
F;497-599/Domain: cadherin repeat homology <CR5>
F;602-708/Domain: cadherin repeat homology <CR6>
F;718-822/Domain: cadherin repeat homology <CR7>
F;831-942/Domain: cadherin repeat homology <CR8>
F;948-1049/Domain: cadherin repeat homology <CR9>
F;1052-1153/Domain: cadherin repeat homology <CR10>
F;1156-1278/Domain: cadherin repeat homology <CR11>
F;1281-1384/Domain: cadherin repeat homology <CR12>
F;1387-1489/Domain: cadherin repeat homology <CR13>
F;1492-1601/Domain: cadherin repeat homology <CR14>
F;1607-1713/Domain: cadherin repeat homology <CR15>
F;1717-1823/Domain: cadherin repeat homology <CR16>
F;1826-1922/Domain: cadherin repeat homology <CR17>
F;1925-2027/Domain: cadherin repeat homology <CR18>
F;2028-2167/Domain: cadherin repeat homology <CR19>
F;2169-2278/Domain: cadherin repeat homology <CR20>
F;2281-2384/Domain: cadherin repeat homology <CR21>
F;2387-2491/Domain: cadherin repeat homology <CR22>
F;2494-2596/Domain: cadherin repeat homology <CR23>
F;2599-2703/Domain: cadherin repeat homology <CR24>
F;2707-2810/Domain: cadherin repeat homology <CR25>
F;2813-2913/Domain: cadherin repeat homology <CR26>
F;2915-3013/Domain: cadherin repeat homology <CR27>
F;3014-3124/Domain: cadherin repeat homology <CR28>
F;3127-3229/Domain: cadherin repeat homology <CR29>
F;3232-3334/Domain: cadherin repeat homology <CR30>
F;3337-3439/Domain: cadherin repeat homology <CR31>
F;3442-3545/Domain: cadherin repeat homology <CR32>
F;3548-3651/Domain: cadherin repeat homology <CR33>
F;3654-3756/Domain: cadherin repeat homology <CR34>
F;3954-4010/Domain: EGF homology <EG1>
F;4017-4048/Domain: EGF homology <EG2>
F;4056-4089/Domain: EGF homology <EG3>
F;4096-4127/Domain: EGF homology <EG4>
F;4584-4609/Domain: transmembrane #status predicted <TM>
F;4610-5147/Domain: intracellular #status predicted <INT>

Query Match 12.7%; Score 76; DB 1; Length 5147;
Best Local Similarity 29.1%; Pred. No. 54;
Matches 25; Conservative 10; Mismatches 41; Indels 10; Gaps 4;
Matches 25; Conservative 10; Mismatches 41; Indels 10; Gaps 4;

Qy 15 SSKPGCYNGKHQYINQWERTYLGVLVCTCYGSGRGNCBKPABETCFDKYTGNT 74
IUFFTM
Db 4011 SRQDPCLPNPCHSOV--QCRR--LGSDFQCMCPANRDGKCEK--ERSDVCYSKPCRN- 4063

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:34:19 ; Search time 55.9041 Seconds
(without alignments)
1090.970 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQMVPQSPVAVSQSKPG.....SMIWDCTCIGAGRGISCTI 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	596	99.3	749	2 Q7Z391	Q7Z391 homo sapien
2	596	99.3	1103	2 Q6MZFA	Q6MZFA homo sapien
3	596	99.3	1103	2 CAE45885	CAE45885 homo sapi
4	596	99.3	1103	2 CAE46200	CAE46200 homo sapi
5	596	99.3	2296	2 Q6N0A6	Q6N0A6 homo sapien
6	596	99.3	2296	2 CAE45714	CAE45714 homo sapi
7	596	99.3	2386	1 F1NC_HUMAN	P02751 homo sapien
8	596	99.3	2444	2 Q6N025	Q6N025 homo sapien
9	596	99.3	2444	2 CAE45847	CAE45847 homo sapi
10	596	99.3	2477	2 Q6MZU5	Q6MZU5 homo sapien
11	596	99.3	2477	2 CAE45932	CAE45932 homo sapi
12	580	96.7	296	2 Q8C6J7	Q8C6J7 mus musculus
13	580	96.7	2477	1 F1NC_MOUSE	P11276 mus musculus
14	575	95.8	2265	1 F1NC_BOVIN	P07589 bos taurus
15	574	95.7	2477	1 F1NC_RAT	P04937 rattus norv
16	492	82.0	2481	1 F1NC_XENLA	Q91740 xenopus lae
17	492	82.0	2481	2 Q6QQA5	Q6QQA5 xenopus lae
18	480	80.0	215	2 Q6DD34	Q6DD34 xenopus lae
19	342	57.0	2408	2 Q6JAN2	Q6JAN2 brachydanio
20	342	57.0	2408	2 AAT08488	AAT08488 brachydan
21	337.5	56.2	922	2 Q93405	Q93405 brachydanio
22	337.5	56.2	2478	2 Q93406	Q93406 brachydanio
23	164	27.3	216	2 Q99KD0	Q99KD0 mus musculus
24	164	27.3	810	2 Q8R3F3	Q8R3F3 mus musculus
25	163.5	27.3	111	2 Q862C9	Q862C9 bos taurus
26	163.5	27.3	412	2 Q71U44	Q71U44 bos taurus
27	163.5	27.3	412	2 AAD33692	AAD33692 bos tauru
28	162.5	27.1	163	2 Q9H382	Q9H382 homo sapien
29	162.5	27.1	211	2 Q7L553	Q7L553 homo sapien
30	162.5	27.1	211	2 AAH05858	AAH05858 homo sapi
31	162.5	27.1	268	2 Q6PJ55	Q6PJ55 homo sapien

32	162.5	27.1	268	2 AAH16875	Aah16875 homo sapi
33	162.5	27.1	1034	2 Q6N084	Q6N084 homo sapien
34	162.5	27.1	1034	2 CAE45786	CAE45786 homo sapi
35	162.5	27.1	1255	2 Q6MZS0	Q6MZS0 homo sapien
36	162.5	27.1	1255	2 CAE45958	CAE45958 homo sapi
37	162.5	27.1	2193	2 Q6MZM7	Q6MZM7 homo sapien
38	162.5	27.1	2193	2 CAE46002	CAE46002 homo sapi
39	161.5	26.9	195	2 Q7T2W7	Q7T2W7 brachydanio
40	161.5	26.9	358	2 Q7TIS2	Q7TIS2 brachydanio
41	151.5	25.2	100	1 F1NC_RABIT	Q28749 oryctolagus
42	148.5	24.8	351	2 O028T6	O028T6 oryctolagus
43	141	23.5	1256	1 F1NC_CHICK	P11722 gallus gall
44	140	23.3	141	2 Q90XQ2	Q90XQ2 ambystoma m
45	140	23.3	320	2 Q95KV4	Q95KV4 bos taurus

ALIGNMENTS

RESULT 1

Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFp686B18150.
GN Name=DKFp686B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothel primary cell culture;
RA Bloecker H., Boecker M., Mewes H.W., Weil B., Amad C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; -.
DR PROSITE; PS00023; FIBRONECTIN 2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 99.3%; Score 596; DB 2; Length 749;

Best Local Similarity 99.1%; Pred. No. 2e-54;

Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYINQQWERTYLGNVLVCTCYGSRGFNCESKP 60
Db 124 QAQMVPQSPVAVSQSKPGCYDNGKHQYINQQWERTYLGNVLVCTCYGSRGFNCESKP 183

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

Db 184 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 229

RESULT 2

Q6MZFA PRELIMINARY; PRT; 1103 AA.
ID Q6MZFA

AC Q6MZFA;

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219)
(Fragment)
GN Name=DKFZp686K139; Synonyms=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG THE GERMAN HUMAN CDNA CONSORTIUM;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -.
DR EMBL; BX640802; CAE45885.1; -.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrnctn1.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRODom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;
Query Match 99.3%; Score 596; DB 2; Length 1103;
Best Local Similarity 99.1%; Pred. No. 3.1e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 QAOQMVQPSVAVSQSPKGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
121 QAOQMVQPSVAVSQSPKGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 180
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
181 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 226
Db
RESULT 3
CAE45885 PRELIMINARY; PRT; 1103 AA.
AC CAE45885
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686F219 (Fragment).
GN DKFZp686F219.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640802; CAE45885.1; -.
KW Hypothetical protein.

FT NON TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;
Query Match 99.3%; Score 596; DB 2; Length 1103;
Best Local Similarity 99.1%; Pred. No. 3.1e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 QAOQMVQPSVAVSQSPKGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
121 QAOQMVQPSVAVSQSPKGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 180
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
181 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 226
Db
RESULT 4
CAE46200 PRELIMINARY; PRT; 1103 AA.
AC CAE46200
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Fragment).
GN DKFZp686K139.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;
Query Match 99.3%; Score 596; DB 2; Length 1103;
Best Local Similarity 99.1%; Pred. No. 3.1e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 QAOQMVQPSVAVSQSPKGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
121 QAOQMVQPSVAVSQSPKGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 180
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
181 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 226
Db
RESULT 5
Q6N0A6 PRELIMINARY; PRT; 2296 AA.
AC Q6N0A6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M04163.
GN Name=DKFZp686M04163;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG THE GERMAN HUMAN CDNA CONSORTIUM;
RA Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; BX640608; CAE45714.1; --
DR InterPro; IPR002086; Aldehyde dehydr.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRODOM; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 15.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 2296 AA; 252761 MW; 9AB2D723CC0CED70 CRC64;

Query Match 99.3%; Score 596; DB 2; Length 2296;
Best Local Similarity 99.1%; Pred. No. 6.8e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAAQMVQPSVAVSQSGPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCESKP 60
DQ 32 QAAQMVQPSVAVSQSGPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCESKP 91
QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCCTCIGAGRGISCTI 106
DQ 92 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCCTCIGAGRGISCTI 137

RESULT 6
CAE45714 PRELIMINARY; PRT; 2296 AA.
ID CAE45714;
AC CAE45714;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M04163.
GN DKFZp686M04163.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RA Bloecher H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640608; CAE45714.1; --
KW Hypothetical protein.
SQ SEQUENCE 2296 AA; 252761 MW; 9AB2D723CC0CED70 CRC64;

Query Match 99.3%; Score 596; DB 2; Length 2296;
Best Local Similarity 99.1%; Pred. No. 6.8e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAAQMVQPSVAVSQSGPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCESKP 60
DQ 32 QAAQMVQPSVAVSQSGPGCYDNGKHQYINQWERTYLGVLVCTCYGSGRGFNCESKP 91
QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCCTCIGAGRGISCTI 106
DQ 92 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCCTCIGAGRGISCTI 137
```

```
RESULT 7
FINC_HUMAN STANDARD; PRT; 2386 AA.
ID FINC_HUMAN STANDARD; PRT; 2386 AA.
AC P02751; O95609; O95610; Q14312; Q14325; Q14326; Q86727; Q8IV18;
AC O96KP7; O96KP8; O96KP9; Q9H1B8; Q9HAP3; Q9UMK2;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Fibronectin precursor (FN) (Cold-insoluble globulin) (CIG).
GN Name=FN1; Synonyms=FN;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21600194; PubMed=11737888;
RA Schor S.L., Schor A.M.;
RT "Phenotypic and genetic alterations in mammary stroma: implications for tumour progression.";
RL Breast Cancer Res. 3:373-379(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 3; 7 AND 10).
RC TISSUE=Cervix;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.-W., Weil B., Amid C., Osanger A., Fobo G., Han M.,
RA Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-38 FROM N.A.
RX MEDLINE=87030890; PubMed=3770189;
RA Gutman A., Yamada K.M., Kornblitt A.R.;
RT "Human fibronectin is synthesized as a pre-polypeptide.";
RL FEBS Lett. 207:145-148(1986).
RN [4]
RP SEQUENCE OF 1-49 FROM N.A.
RX MEDLINE=87175578; PubMed=3031656;
RA Dean D.C., Bowlus C.L., Bourgeois S.;
RT "Cloning and analysis of the promoter region of the human fibronectin gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:1876-1880(1987).
RN [5]
RP SEQUENCE OF 28-2386 FROM N.A. (ISOFORM 3).
RX MEDLINE=85284965; PubMed=2992939;
RA Kornblitt A.R., Umezawa K., Vibe-Pedersen K., Baralle F.E.;
RT "Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene.";
RL EMBO J. 4:1755-1759(1985).
RN [6]
RP SEQUENCE OF 103-481 AND 2228-2386 FROM N.A. (ISOFORMS 1; 3; 8 AND 9).
RC TISSUE=Periphereal blood T-cell, and Umbilical vein endothelial cells;
RA Godfrey H.P., Ebrahim A.A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 973-2386 FROM N.A. (ISOFORM 3).
RX MEDLINE=84272258; PubMed=6462919;
RA Kornblitt A.R., Vibe-Pedersen K., Baralle F.E.;
RT "Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats.";
RL Nucleic Acids Res. 12:5853-5868(1984).
RN [8]
RP SEQUENCE OF 1232-1782 FROM N.A. (ISOFORM 7).
RX MEDLINE=8823940; PubMed=3375063;
RA Paoletta G., Henschliffe C., Sebastio G., Baralle F.E.;
RT "Sequence analysis and in vivo expression show that alternative splicing of ED-B and ED-A regions of the human fibronectin gene are independent events.";
RL Nucleic Acids Res. 16:3545-3557(1988).
RN [9]
RP SEQUENCE OF 1257-1365 FROM N.A. (ISOFORM 11).
RX MEDLINE=88041070; PubMed=3478690;
RA Gutman A., Kornblitt A.R.;
RT "Identification of a third region of cell-specific alternative
```


RT splicing in human fibronectin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7179-7182(1987).
RN [10]
RP SEQUENCE OF 1441-1548.
RX MEDLINE=82265604; PubMed=7050098;
RA Pierschbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;
RT "The cell attachment domain of fibronectin. Determination of the
RT primary structure.";
RL J. Biol. Chem. 257:9593-9597(1982).
RN [11]
RP SEQUENCE OF 1448-1540 FROM N.A.
RX MEDLINE=83290929; PubMed=6688418;
RA Oldberg A., Linney E., Ruoslahti E.;
RT "Molecular cloning and nucleotide sequence of a cDNA clone coding for
RT the cell attachment domain in human fibronectin.";
RL J. Biol. Chem. 258:10193-10196(1983).
RN [12]
RP SEQUENCE OF 1448-1540 FROM N.A.
RX MEDLINE=86111901; PubMed=3003095;
RA Oldberg A., Ruoslahti E.;
RT "Evolution of the fibronectin gene. Exon structure of cell attachment
RT domain.";
RL J. Biol. Chem. 261:2113-2116(1986).
RN [13]
RP SEQUENCE OF 1594-2386 FROM N.A. (ISOFORM 1).
RX MEDLINE=85280409; PubMed=292573;
RA Bernard M.P., Kolbe M., Weil D., Chu M.-L.;
RT "Human cellular fibronectin: comparison of the carboxyl-terminal
RT portion with rat identifies primary structural domains separated by
RT hypervariable regions.";
RL Biochemistry 24:2698-2704(1985).
RN [14]
RP SEQUENCE OF 1712-1739 FROM N.A.
RX MEDLINE=87026578; PubMed=3021206;
RA Sekiguchi K., Klos A.M., Kurachi K., Yoshitake S., Hakomori S.;
RT "Human liver fibronectin complementary DNAs: identification of two
RT different messenger RNAs possibly encoding the alpha and beta subunits
RT of plasma fibronectin.";
RL Biochemistry 25:4936-4941(1986).
RN [15]
RP SEQUENCE OF 1788-2386 FROM N.A. (ISOFORMS 4; 5 AND 6).
RC TISSUE=Cartilage; PubMed=12127832;
RX MEDLINE=22126816; PubMed=12127832;
RA Parker A.E., Boutell J., Carr A., Maciewicz R.A.;
RT "Novel cartilage-specific splice variants of fibronectin.";
RL Osteoarthritis Cartilage 10:528-534(2002).
RN [16]
RP SEQUENCE OF 32-290.
RX MEDLINE=84032463; PubMed=6630202;
RA Garcia-Pardo A., Pearlstein E., Frangione B.;
RT "Primary structure of human plasma fibronectin. The 29,000-dalton NH2-
RT terminal domain.";
RL J. Biol. Chem. 258:12670-12674(1983).
RN [17]
RP SEQUENCE OF 309-608, AND COLLAGEN-BINDING.
RX MEDLINE=87080265; PubMed=3024962;
RA Owens R.J., Baralle F.E.;
RT "Mapping the collagen-binding site of human fibronectin by expression
RT in Escherichia coli.";
RL EMBO J. 5:2825-2830(1986).
RN [18]
RP SULFATION.
RX MEDLINE=86042625; PubMed=2414772;
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
RT "Tyrosine sulfation of proteins from the human hepatoma cell line
RT HepG2.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
RN [19]
RP O-GLYCOSYLATION OF THR-2064.
RX MEDLINE=9119085; PubMed=2012601;
RA Tressel T., McCarthy J.B., Calaycay J., Lee T.D., Legesse K.,
RA Shively J.E., Pande H.;
RT "Human plasma fibronectin. Demonstration of structural differences

RT between the A- and B-chains in the III CS region.";
RL Biochem. J. 274:731-738(1991).
RN [20]
RP FBLNI-BINDING SITE.
RX MEDLINE=93015879; PubMed=1400330;
RA Balbona K., Tran H., Godyna S., Strickland D.K.,
RA Argraves W.S.;
RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding
RT region of fibronectin.";
RL J. Biol. Chem. 267:20120-20125(1992).
RN [21]
RP CHARACTERIZATION OF FIBRIN-BINDING SITE 1.
RX MEDLINE=95081153; PubMed=7989369;
RA Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;
RT "Further characterization of the NH2-terminal fibrin-binding site on
RT fibronectin.";
RL J. Biol. Chem. 269:31938-31945(1994).
RN [22]
RP INTERACTION WITH LGALS3BP.
RX PubMed=9501082;
RA Sasaki T., Brakebusch C., Engel J., Timpl R.;
RT "Mac-2 binding protein is a cell-adhesive protein of the extracellular
RT matrix which self-assembles into ring-like structures and binds beta1
RT integrins, collagens and fibronectin.";
RL EMBO J. 17:1606-1613(1998).
RN [23]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE=92162710; PubMed=13111202;
RA Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,
RA Campbell I.D.;
RT "1H NMR assignment and secondary structure of the cell adhesion type
RT III module of fibronectin.";
RL Biochemistry 31:2068-2073(1992).
RN [24]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE=93046665; PubMed=1423622;
RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;
RT "The three-dimensional structure of the tenth type III module of
RT fibronectin: an insight into RGD-mediated interactions.";
RL Cell 71:671-678(1992).
RN [25]
RP STRUCTURE BY NMR OF 182-275.
RX MEDLINE=94141923; PubMed=8308892;
RA Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,
RA Campbell I.D.;
RT "Solution structure of a pair of fibronectin type 1 modules with
RT fibrin binding activity.";
RL J. Mol. Biol. 235:1302-1311(1994).
RN [26]
RP STRUCTURE BY NMR OF 32-92.
RX MEDLINE=96069779; PubMed=7583666;
RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;
RT "High-resolution structural studies of the factor XIIIa crosslinking
RT site and the first type 1 module of fibronectin.";
RL Nat. Struct. Biol. 2:946-950(1995).
RN [27]

Query Match 99.3%; Score 596; DB 1; Length 2386;
Best Local Similarity 99.1%; Pred. No. 7e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAOQMVPQSPVAVSOSKPGCYDNGKHQINQOWERTYLGNAVLCVTCYSGRGFNCESKP 60
DB 32 QAOQMVPQSPVAVSOSKPGCYDNGKHQINQOWERTYLGNAVLCVTCYSGRGFNCESKP 91
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
DB 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 137

RESULT 8
Q6N025 PRELIMINARY; PPT; 2444 AA.
ID Q6N025

AC Q6N025;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M2451 (Fragment).
GN Name=DKFZp686M2451;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RG THE GERMAN HUMAN CDNA CONSORTIUM;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640731; CAE45847.1; -.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrctnl.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR00562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 16.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 16.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 16.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2444 AA; 268676 MW; 71C5E8C56A84C7BC CRC64;

Query Match 99.3%; Score 596; DB 2; Length 2444;
Best Local Similarity 99.1%; Pred. No. 7.2e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
DB 121 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 180

QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
DB 181 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 226

RESULT 9
CAE45847 PRELIMINARY; PRT; 2444 AA.
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M2451 (Fragment).
GN DKFZp686M2451.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

Query Match 99.3%; Score 596; DB 2; Length 2444;
Best Local Similarity 99.1%; Pred. No. 7.2e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
DB 121 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 180

QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
DB 181 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 226

RESULT 9
CAE45847 PRELIMINARY; PRT; 2444 AA.
DT 02-MAR-2004 (TReMBLrel. 27, Created)
DT 02-MAR-2004 (TReMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M2451 (Fragment).
GN DKFZp686M2451.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX640731; CAE45847.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2444 AA; 268676 MW; 71C5E8C56A84C7BC CRC64;

Query Match 99.3%; Score 596; DB 2; Length 2444;
Best Local Similarity 99.1%; Pred. No. 7.2e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
DB 121 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 180

QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
DB 181 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 226

RESULT 10
Q6MZU5 PRELIMINARY; PRT; 2477 AA.
AC Q6MZU5;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein DKFZp686O1166.
GN Name=DKFZp686O1166;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RG THE GERMAN HUMAN CDNA CONSORTIUM;
RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640875; CAE45932.1; -.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrctnl.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR00562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 17.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 17.
KW Hypothetical protein.
SQ SEQUENCE 2477 AA; 272335 MW; D358D85C6B18207C CRC64;

Query Match 99.3%; Score 596; DB 2; Length 2477;
Best Local Similarity 99.1%; Pred. No. 7.3e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 60
DB 32 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCESKP 91

QY 61 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
DB 92 EAETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 137

[illegible]

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-920 FROM N.A.
RC STRAIN=FVB/N-3; TISSUE=Breast tumor;
RX MEDLINE=23288357; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 1-28 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94131313; PubMed=8299972;
RA Polly P., Nicholson R.C.;
RT "Sequence of the mouse fibronectin-encoding gene promoter region.";
RL Gene 137:353-354(1993).
RN [3]
RP SEQUENCE OF 562-834 FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=95403556; PubMed=7673336;
RA Tait J.F., Weller A., Timpl R., Ekblom M., Ekblom P.;
RT "Regulation of mesenchymal extracellular matrix protein synthesis by
transforming growth factor-beta and glucocorticoids in tumor stroma.";
RL J. Cell Sci. 108:2153-2162(1995).
RN [4]
RP SEQUENCE OF 899-2376 FROM N.A.
RA Gorski G., Aros M., Norton P.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 2375-2477 FROM N.A.
RA Blatti S.P., Foster D.N., Ranganathan G., Moses H.L., Getz M.J.;
RX MEDLINE=88124987; PubMed=3124113;
RT "Induction of fibronectin gene transcription and mRNA is a primary
response to growth-factor stimulation of AKR-2B cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1119-1123(1988).
RN [6]
RP SEQUENCE OF 2375-2477 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=93011702; PubMed=1327855;
RA Khandjian E.W., Salomon C., Leonard N., Tremblay S., Turler H.;
RT "Fibronectin gene expression in proliferating, quiescent, and SV40-
infected mouse kidney cells.";
RL Exp. Cell Res. 202:464-470(1992).
RN [7]
RP STRUCTURE BY NMR OF 1447-1630.
RX MEDLINE=98202578; PubMed=9533887;
RA Copie V., Tomita Y., Akiyama S.K., Aota S., Yamada K.M., Venable R.M.,
RA Pastor R.W., Krueger S., Torchia D.A.;
RT "Solution structure and dynamics of linked cell attachment modules of
mouse fibronectin containing the RGD and synergy regions: comparison
with the human fibronectin crystal structure.";
RL J. Mol. Biol. 277:663-682(1998).
RN [8]
RP DOWN-REGULATION BY GLUCOCORTICOID.
RX MEDLINE=21600963; PubMed=11737251;
RA Gu Y.-C., Tait J.F., Gullberg D., Timpl R., Ekblom M.;
RT "Glucocorticoids down-regulate the extracellular matrix proteins
fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma.";
RL Eur. J. Haematol. 67:176-184(2001).
CC -i- FUNCTION: Fibronectins bind cell surfaces and various compounds
including collagen, fibrin, heparin, DNA, and actin. Fibronectins
are involved in cell adhesion, cell motility, opsonization, wound
healing, and maintenance of cell shape.
CC -i- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
variants, connected by 2 disulfide bonds near the carboxyl ends;
to a lesser extent homodimers. Interacts with FBLN1 and LGALS3BP
(by similarity).
CC -i- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -i- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=1;
Comment=A number of isoforms are produced. Each of the "extra
domain" and the connecting strand 3 are present in some forms of
fibronectin and absent in others;
CC Name=1;
CC IsoId=SP11276-1; Sequence=Displayed;
CC -i- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
form), made by fibroblasts, epithelial and other cell types, is
deposited as fibrils in the extracellular matrix.
CC -i- INDUCTION: Glucocorticoids suppressed mRNA expression and protein
synthesis.
CC -i- PTM: Sulfated (by similarity).
CC -i- SIMILARITY: Contains 12 fibronectin type I domains.
CC -i- SIMILARITY: Contains 2 fibronectin type II domains.
CC -i- SIMILARITY: Contains 17 fibronectin type III domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; BC051082; AAH51082.1; -
DR EMBL; Z22729; CAA80422.1; -
DR EMBL; X82402; CAA57796.1; -
DR EMBL; X93167; CAA63654.1; -
DR EMBL; M18194; AAA37636.1; -
DR EMBL; S45680; AAB23491.1; -
DR FIR; A49173; A49173.
DR FIR; I48349; I48349.
DR PDB; 1MFN; NMR; @=1446-1630.
DR PDB; 2MFN; NMR; @=1446-1630.
DR MGD; MGI:95566; Fnl.
DR GO; GO:0007155; P:cell adhesion; IDA.
DR GO; GO:0007044; P:cell-substrate junction assembly; IDA.
DR GO; GO:0042060; P:wound healing; IMP.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrctn1.
DR InterPro; IPR003961; FN_III subd.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00014; FNTYPEIII.
DR ProDom; PD000995; FN_Type_II; 2.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 17.
KW 3D-structure; Acute phase; Alternative splicing; Cell adhesion;
KW Glycoprotein; Heparin-binding; Phosphorylation; Plasma; Repeat;
KW Signal; Sulfation.
FT SIGNAL 1 32 By similarity.
FT CHAIN 33 2477 Fibronectin.
FT DOMAIN 53 273 Fibrin- and heparin-binding 1.

```
FT DOMAIN 308 608
FT DNA_BIND 906 1171
FT 1357 1630
FT 1811 2081
FT 2296 2427
FT 51 96
FT 96 140
FT 140 185
FT 185 230
FT 230 272
FT 306 343
FT 343 404
FT 404 469
FT 469 516
FT 516 559
FT 559 602
FT 602 699
FT 699 719
FT 719 808
FT 808 810
FT 810 897
FT 897 905
FT 905 995
FT 995 1091
FT 1091 1171
FT 1171 1264
FT 1264 1355
FT 1355 1446
FT 1446 1536
FT 1536 1631
FT 1631 1720
FT 1720 1810
FT 1810 1900
FT 1900 1993
FT 1993 2081
FT 2081 2201
FT 2201 2280
FT 2280 2338
FT 2338 2381
FT 2381 2426
FT 2426 2614
FT 2614 2813
FT 2813 53
FT 53 79
FT 79 88
FT 88 126
FT 126 136
FT 136 142
FT 142 170
FT 170 168

Collagen-binding.
CELL-ATTACHMENT.
Heparin-binding 2.
Fibrin-binding 2.
Fibronectin type-I 1.
Fibronectin type-I 2.
Fibronectin type-I 3.
Fibronectin type-I 4.
Fibronectin type-I 5.
Fibronectin type-I 6.
Fibronectin type-II 1.
Fibronectin type-II 2.
Fibronectin type-I 7.
Fibronectin type-I 8.
Fibronectin type-I 9.
Fibronectin type-III 1.
Fibronectin type-III 2.
Fibronectin type-III 3.
Fibronectin type-III 4.
Fibronectin type-III 5.
Fibronectin type-III 6.
Fibronectin type-III 7.
Fibronectin type-III 8 (extra domain 1).
Fibronectin type-III 9.
Fibronectin type-III 10.
Fibronectin type-III 11.
Fibronectin type-III 12.
Fibronectin type-III 13 (extra domain 2).
Fibronectin type-III 14.
Fibronectin type-III 15.
Fibronectin type-III 16.
CONNECTING STRAND 3 (CS-3) (V REGION).
Fibronectin type-III 17.
Fibronectin type-I 10.
Fibronectin type-I 11.
Fibronectin type-I 12.
Cell attachment site.
Cell attachment site.
By similarity.
By similarity.
By similarity.
By similarity.
By similarity.
By similarity.

Query Match 96.7%; Score 580; DB 1; Length 2477;
Best Local Similarity 95.3%; Pred. No. 3 6e-52;
Matches 101; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 1 QAAQMVQPSVAVSQSGKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGFNCSKP 60
Db 33 QAAQIVQPSVAVSQSGKPGCYDNGKHQYQINQWERTYLGVALVCTCYGSGRGFNCSKP 92

Oy 61 EAETCFDKYGTNTVRGDTYRPRKDSMIWDCTCIGAGRGRI SCTI 106
Db 93 EPETCFDKYGTNTYKVGDTYRPRKDSMIWDCTCIGAGRGRI SCTI 138

RESULT 14
FIND_BOVIN
ID_FIND_BOVIN STANDARD; PRT; 2265 AA.
AC P07589;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fibronectin (FN).
GN Name=FN1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos
OX NCBI_TaxID=9913;
```

KW Acute phase; Alternative splicing; Cell adhesion; Cell shape;
KW Direct protein sequencing; Glycoprotein; Heparin-binding;
KW Phosphorylation; Plasma; Pyrrolidone carboxylic acid; Repeat;
KW Sulfation.

FT MOD_RES 1 1 Pyrrolidone carboxylic acid.
FT DOMAIN 21 241 Fibrin- and heparin-binding 1.
FT DOMAIN 277 577 Collagen-binding.
FT DNA_BIND 876 1141
FT DOMAIN 1236 1509 Heparin-binding 2.
FT DOMAIN 1600 1870 Fibrin-binding 2.
FT DOMAIN 1991 2216 Fibrinectin type-I 1.
FT DOMAIN 19 59 Fibrinectin type-I 2.
FT DOMAIN 64 107 Fibrinectin type-I 3.
FT DOMAIN 108 151 Fibrinectin type-I 4.
FT DOMAIN 153 197 Fibrinectin type-I 5.
FT DOMAIN 198 242 Fibrinectin type-I 6.
FT DOMAIN 275 314 Fibrinectin type-II 1.
FT DOMAIN 314 373 Fibrinectin type-II 2.
FT DOMAIN 374 438 Fibrinectin type-I 7.
FT DOMAIN 437 480 Fibrinectin type-I 8.
FT DOMAIN 485 527 Fibrinectin type-I 9.
FT DOMAIN 528 571 Fibrinectin type-III 1.
FT DOMAIN 576 668 Fibrinectin type-III 2.
FT DOMAIN 689 778 Fibrinectin type-III 3.
FT DOMAIN 780 867 Fibrinectin type-III 4.
FT DOMAIN 877 964 Fibrinectin type-III 5.
FT DOMAIN 965 1053 Fibrinectin type-III 6.
FT DOMAIN 1056 1141 Fibrinectin type-III 7.
FT DOMAIN 1142 1234 Fibrinectin type-III 8.
FT DOMAIN 1235 1325 Fibrinectin type-III 9.
FT DOMAIN 1326 1415 Fibrinectin type-III 10.
FT DOMAIN 1416 1505 Fibrinectin type-III 11 (extra domain).
FT DOMAIN 1510 1599 Fibrinectin type-III 12.
FT DOMAIN 1602 1689 Fibrinectin type-III 13.
FT DOMAIN 1692 1780 Fibrinectin type-III 14.
FT DOMAIN 1781 1870 CONNECTING STRAND 3 (CS-3) (V REGION).
FT DOMAIN 1871 1930 Fibrinectin type-III 15.
FT DOMAIN 1931 1979 Fibrinectin type-I 1.
FT DOMAIN 2083 2127 Fibrinectin type-I 11.
FT DOMAIN 2128 2170 Fibrinectin type-I 12.
FT DOMAIN 2172 2215 Cell attachment site.
FT SITE 1493 1495

FT DISULFID 21 47
FT DISULFID 45 56
FT DISULFID 66 94
FT DISULFID 92 104
FT DISULFID 110 138
FT DISULFID 136 148
FT DISULFID 155 184
FT DISULFID 182 194
FT DISULFID 200 229
FT DISULFID 227 239
FT DISULFID 277 304
FT DISULFID 302 311
FT DISULFID 329 355
FT DISULFID 343 370
FT DISULFID 389 415
FT DISULFID 403 430
FT DISULFID 439 467
FT DISULFID 465 477
FT DISULFID 487 514
FT DISULFID 512 524
FT DISULFID 530 558
FT DISULFID 556 568
FT DISULFID 2085 2114
FT DISULFID 2112 2124
FT DISULFID 2130 2157
FT DISULFID 2155 2167
FT DISULFID 2174 2200
FT DISULFID 2198 2209
FT DISULFID 2246 2246
FT DISULFID 2250 2250

FT MOD_RES 845
FT MOD_RES 845 Interchain (with C-2250).
FT MOD_RES 845 Interchain (with C-2246).
FT MOD_RES 845 Sulfotyrosine (Potential).

FT MOD_RES 850 850 Sulfotyrosine (Potential).
FT CARBOHYD 399 399 N-linked (GlcNAc...)
FT CARBOHYD 497 497 N-linked (GlcNAc...)
FT CARBOHYD 511 511 N-linked (GlcNAc...)
FT CARBOHYD 846 846 N-linked (GlcNAc...)
FT CARBOHYD 976 976 N-linked (GlcNAc...)
FT CARBOHYD 1213 1213 N-linked (GlcNAc...)
FT CARBOHYD 1987 1987 N-linked (GlcNAc...)
FT CARBOHYD 1943 1943 O-linked (GalNAc...)
FT CARBOHYD 1944 1944 O-linked (GalNAc...)
FT MOD_RES 2263 2263 Phosphoserine.
SQ SEQUENCE 2265 AA; 249557 MW; C2D21D486F498D5C CRC64;

Query Match 95.8%; Score 575; DB 1; Length 2265;
Best Local Similarity 94.3%; Pred. No. 1.1e-51;
Matches 100; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 QAOQMVPQSPVAVSQSKPCYDNGKHYQINQOQWERTYLGVLVCTCYGSGRGNCSKP 60
DB 1 QAOQIVQPQSPVAVSQSKPCYDNGKHYQINQOQWERTYLGVLVCTCYGSGRGNCSKP 60
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
DB 61 EPBETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

RESULT 15
FIND_RAT STANDARD; PRT; 2477 AA.
ID_FIND_RAT AC P04937;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fibrinectin precursor (FN).
GN Name=fnl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=88054951; PubMed=2445560;
RA Schwarzbauer J.E., Patel R.S., Fonda D., Hynes R.O.;
RT "Multiple sites of alternative splicing of the rat fibrinectin gene transcript.";
RL EMBO J. 6:2573-2580 (1987).
RN [2]
RP SEQUENCE OF 1-139 AND 2382-2477 FROM N.A.
RC STRAIN=Fischer; TISSUE=Liver;
RX MEDLINE=88054950; PubMed=3119323;
RA Patel R.S., Odermatt E., Schwarzbauer J.E., Hynes R.O.;
RT "Organization of the fibrinectin gene provides evidence for exon shuffling during evolution.";
RL EMBO J. 6:2565-2572 (1987).
RN [3]
RP SEQUENCE OF 1586-2477 FROM N.A.
RX MEDLINE=84082067; PubMed=6317187;
RA Schwarzbauer J.E., Tamkun J.W., Lemischka I.R., Hynes R.O.;
RT "Three different fibrinectin mRNAs arise by alternative splicing within the coding region.";
RL Cell 35:421-431 (1983).
CC -!- FUNCTION: Fibrinectins bind cell surfaces and various compounds including collagen, fibrin, heparin, DNA, and actin. Fibrinectins are involved in cell adhesion, cell motility, opsonization, wound healing, and maintenance of cell shape.
CC -!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced variants, connected by 2 disulfide bonds near the carboxyl ends; to a lesser extent homodimers. Interacts with FBLN1 and LGALS3BP (by similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;

CC Comment=Each of the "extra domain" and the connecting strand 3
 CC are present in some forms of fibronectin and absent in others;
 CC Name=1;
 CC IsoId=P04937-1; Sequence=Displayed;
 CC Name=2; Synonyms=FNIII-13-less;
 CC IsoId=P04937-2; Sequence=VSP_003258;
 CC Name=3; Synonyms=Lambda-RLF4-5;
 CC IsoId=P04937-3; Sequence=VSP_003259;
 CC Name=4; Synonyms=Lambda-RLF6;
 CC IsoId=P04937-4; Sequence=VSP_003260;
 CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
 CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
 CC forms), made by fibroblasts, epithelial and other cell types, is
 CC deposited as fibrils in the extracellular matrix.
 CC -!- PTM: Sulfated (By similarity)
 CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
 CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
 CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X15906; CAA34020.1; -
 CC EMBL; L29191; AAA41166.1; -
 CC EMBL; L00191; AAA41166.1; JOINED.
 CC EMBL; L29191; AAA41167.1; -
 CC EMBL; L00191; AAA41167.1; JOINED.
 CC EMBL; L29191; AAA41168.1; -
 CC EMBL; L00191; AAA41168.1; JOINED.
 CC EMBL; X05831; CRA29278.1; -
 CC EMBL; X05832; CRA29279.1; -
 CC EMBL; X05833; CRA29280.1; -
 CC EMBL; X05834; CRA29281.1; -
 CC PIR; S14428; S14428.
 CC HSSP; P08253; IKS0.
 CC RGD; 2624; Fnl.
 CC InterPro; IPR006209; EGF like.
 CC InterPro; IPR000083; Fibrnctnl.
 CC InterPro; IPR003962; FNIII subd.
 CC InterPro; IPR003961; FNIII.
 CC InterPro; IPR008957; FNIII-like.
 CC InterPro; IPR000562; FN_Type_II.
 CC Pfam; PF00039; fn1; 12.
 CC Pfam; PF00040; fn2; 2.
 CC Pfam; PF00041; fn3; 17.
 CC PRINTS; PR00013; FNTYPEII.
 CC PRINTS; PR00014; FNTYPEIII.
 CC ProDom; PD000995; FN_Type_II; 2.
 CC PROSITE; PS00022; EGF 1; 2.
 CC PROSITE; PS01253; FIBRONECTIN 1; 12.
 CC PROSITE; PS00023; FIBRONECTIN_2; 2.
 CC PROSITE; PS00853; FN3; 17.
 CC Acute phase; Alternative splicing; Cell adhesion; Glycoprotein;
 CC Heparin-binding; Phosphorylation; Plasma; Repeat; Signal; Sulfation.
 CC SIGNAL 1 32
 CC CHAIN 33 2477 Fibronectin.
 CC DOMAIN 53 273 Fibrin- and heparin-binding 1.
 CC DOMAIN 308 608 Collagen-binding.
 CC DNA BIND 906 1171
 CC DOMAIN 1357 1630 CELL-ATTACHMENT.
 CC DOMAIN 1811 2081 Heparin-binding 2.
 CC DOMAIN 2296 2427 Fibrin-binding 2.
 CC DOMAIN 51 91 Fibronectin type-I 1.
 CC DOMAIN 96 139 Fibronectin type-I 2.
 CC DOMAIN 140 183 Fibronectin type-I 3.
 CC DOMAIN 185 229 Fibronectin type-I 4.
 CC DOMAIN 230 274 Fibronectin type-I 5.
 CC DOMAIN 306 345 Fibronectin type-I 6.

FT	DOMAIN	345	404	Fibronectin type-II 1.
FT	DOMAIN	405	469	Fibronectin type-II 2.
FT	DOMAIN	468	511	Fibronectin type-I 7.
FT	DOMAIN	516	558	Fibronectin type-I 8.
FT	DOMAIN	559	602	Fibronectin type-I 9.
FT	DOMAIN	607	699	Fibronectin type-III 1.
FT	DOMAIN	719	808	Fibronectin type-III 2.
FT	DOMAIN	810	897	Fibronectin type-III 3.
FT	DOMAIN	907	994	Fibronectin type-III 4.
FT	DOMAIN	995	1083	Fibronectin type-III 5.
FT	DOMAIN	1091	1171	Fibronectin type-III 6.
FT	DOMAIN	1172	1264	Fibronectin type-III 7.
FT	DOMAIN	1265	1355	Fibronectin type-III 8 (extra domain 1).
FT	DOMAIN	1356	1446	Fibronectin type-III 9.
FT	DOMAIN	1447	1536	Fibronectin type-III 10.
FT	DOMAIN	1537	1626	Fibronectin type-III 11.
FT	DOMAIN	1631	1720	Fibronectin type-III 12.
FT	DOMAIN	1721	1810	Fibronectin type-III 13 (extra domain 2).
FT	DOMAIN	1813	1900	Fibronectin type-III 14.
FT	DOMAIN	1903	1991	Fibronectin type-III 15.
FT	DOMAIN	1992	2081	Fibronectin type-III 16.
FT	DOMAIN	2082	2201	CONNECTING STRAND 3 (CS-3) (V REGION).
FT	DOMAIN	2190	2280	Fibronectin type-III 17.
FT	DOMAIN	2294	2338	Fibronectin type-I 10.
FT	DOMAIN	2339	2381	Fibronectin type-I 11.
FT	DOMAIN	2383	2426	Fibronectin type-I 12.
FT	SITE	1614	1616	Cell attachment site.
FT	SITE	2181	2183	Cell attachment site.
FT	DISULFID	53	79	By similarity.
FT	DISULFID	77	88	By similarity.
FT	DISULFID	98	126	By similarity.
FT	DISULFID	124	136	By similarity.
FT	DISULFID	142	170	By similarity.
FT	DISULFID	168	180	By similarity.
FT	DISULFID	187	216	By similarity.
FT	DISULFID	214	226	By similarity.
FT	DISULFID	232	261	By similarity.
FT	DISULFID	259	271	By similarity.
FT	DISULFID	308	335	By similarity.
FT	DISULFID	333	342	By similarity.
FT	DISULFID	360	386	By similarity.
FT	DISULFID	374	401	By similarity.
FT	DISULFID	420	446	By similarity.
FT	DISULFID	434	461	By similarity.
FT	DISULFID	470	498	By similarity.
FT	DISULFID	496	508	By similarity.
FT	DISULFID	518	545	By similarity.
FT	DISULFID	543	555	By similarity.
FT	DISULFID	561	589	By similarity.
FT	DISULFID	587	599	By similarity.
FT	DISULFID	2296	2325	By similarity.
FT	DISULFID	2323	2335	By similarity.
FT	DISULFID	2341	2368	By similarity.
FT	DISULFID	2366	2378	By similarity.
FT	DISULFID	2385	2409	By similarity.
FT	DISULFID	2407	2423	By similarity.
FT	DISULFID	2458	2458	Interchain (with C-2462).
FT	DISULFID	2462	2462	Interchain (with C-2458).
FT	MOD_RES	875	875	Sulfoxyrosine (Potential).
FT	MOD_RES	880	880	Sulfoxyrosine (Potential).
FT	MOD_RES	2392	2392	Sulfoxyrosine (Potential).
FT	MOD_RES	2475	2475	Phosphoserine (By similarity).
FT	CARBOHYD	430	430	N-linked (GlcNAc...) (By similarity).
FT	CARBOHYD	528	528	N-linked (GlcNAc...) (By similarity).
FT	CARBOHYD	542	542	N-linked (GlcNAc...) (By similarity).
FT	CARBOHYD	876	876	N-linked (GlcNAc...) (Potential).
FT	CARBOHYD	1006	1006	N-linked (GlcNAc...) (Potential).
FT	CARBOHYD	1243	1243	N-linked (GlcNAc...) (By similarity).
FT	CARBOHYD	2198	2198	N-linked (GlcNAc...) (By similarity).
FT	CARBOHYD	2154	2154	O-linked (GlcNAc...) (By similarity).
FT	VARSPPLIC	1720	1809	Missing (in isoform 2).
FT	VARSPPLIC			/FTid=VSP_003258.
FT	VARSPPLIC	2082	2106	Missing (in isoform 3).

```
FT VARSPLIC 2082 2200 /FTId=VSP 003259.
FT Missing (in isoform 4) .
FT /FTId=VSP 003260.
FT CONFLICT 2318 2318 G -> A (in Ref. 3).
SQ SEQUENCE 2477 AA; 272510 MW; B4391A472ECEDEB5 CRC64;

Query Match 95.7%; Score 574; DB 1; Length 2477;
Best Local Similarity 94.3%; Pred. No. 1.6e-51;
Matches 100; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QAAQMVQPOS PVAVSQSKGCGYDNGKHQYQINQOWERTYLGNI VLVCTCYGSGRGFNCESKP 60
Db 33 QAAQIVQPPSPVAVSQSKGCGFDNGKHQYQINQOWERTYLGNI VLVCTCYGSGRGFNCESKP 92
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTI 106
Db 93 EPEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTI 138
```

Search completed: November 3, 2004, 23:54:37
Job time : 56.9041 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:42:25 ; Search time 12.887 Seconds
(without alignments)
545.489 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600
Sequence: 1 QAQMVPQSPVAVSQSKPG.....SMWDCTCIGAGRGRICTI 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	600	100.0	2231	1 US-08-153-799-16	Sequence 16, Appl
2	600	100.0	2324	1 US-08-283-857-1	Sequence 1, Appli
3	600	100.0	2324	5 PCT-US95-09819-1	Sequence 1, Appli
4	600	100.0	2327	6 5455158-1	Patent No. 5455158
5	600	100.0	2386	2 US-09-016-366A-12	Sequence 12, Appl
6	600	100.0	2446	2 US-08-551-356-2	Sequence 2, Appli
7	600	100.0	2446	5 PCT-US93-12687-2	Sequence 2, Appli
8	596	99.3	2386	4 US-09-961-403-1	Sequence 1, Appli
9	162.5	27.1	188	1 US-08-142-449B-14	Sequence 14, Appl
10	162.5	27.1	422	2 US-08-836-854-12	Sequence 12, Appl
11	162.5	27.1	474	2 US-08-836-854-9	Sequence 9, Appli
12	100	16.7	42	1 US-08-340-812-7	Sequence 7, Appli
13	100	16.7	42	1 US-08-459-064B-7	Sequence 7, Appli
14	100	16.7	42	2 US-08-460-421A-7	Sequence 7, Appli
15	100	16.7	42	5 PCT-US93-00909-7	Sequence 7, Appli
16	80.5	13.4	43	2 US-08-982-597A-21	Sequence 21, Appl
17	80.5	13.4	43	3 US-09-136-218-21	Sequence 21, Appl
18	80	13.3	48	2 US-08-982-597A-22	Sequence 22, Appl
19	80	13.3	48	3 US-09-136-218-22	Sequence 22, Appl
20	79.5	13.2	2523	3 US-08-185-432-18	Sequence 18, Appl
21	79.5	13.2	2523	4 US-08-899-232-3	Sequence 3, Appli
22	79.5	13.2	2523	4 US-09-121-457-3	Sequence 3, Appli
23	78.5	13.1	48	2 US-08-982-597A-20	Sequence 20, Appl
24	78.5	13.1	48	3 US-09-136-218-20	Sequence 20, Appl
25	76.5	12.8	1193	2 US-08-400-159-10	Sequence 10, Appl
26	76.5	12.8	1193	3 US-08-611-729A-10	Sequence 10, Appl
27	76.5	12.8	1193	4 US-09-195-524-10	Sequence 10, Appl

28	76	12.7	2556	1 US-08-185-432-17	Sequence 17, Appl
29	76	12.7	2556	4 US-08-899-232-2	Sequence 2, Appli
30	76	12.7	2556	4 US-09-121-457-2	Sequence 2, Appli
31	75.5	12.6	1055	3 US-09-214-278-2	Sequence 2, Appli
32	75.5	12.6	1055	4 US-09-855-722-2	Sequence 2, Appli
33	75.5	12.6	1055	2 US-08-400-159-8	Sequence 8, Appli
34	75.5	12.6	1148	3 US-08-882-046-4	Sequence 4, Appli
35	75.5	12.6	1148	4 US-09-566-047-4	Sequence 4, Appli
36	75.5	12.6	1212	3 US-09-214-278-3	Sequence 3, Appli
37	75.5	12.6	1212	4 US-09-855-722-3	Sequence 3, Appli
38	75.5	12.6	1257	3 US-08-611-729A-8	Sequence 8, Appli
39	75.5	12.6	1257	4 US-09-195-524-8	Sequence 8, Appli
40	75	12.5	950	4 US-10-009-332-1	Sequence 1, Appli
41	74.5	12.4	1238	3 US-09-214-278-5	Sequence 5, Appli
42	74.5	12.4	1238	4 US-09-855-722-5	Sequence 5, Appli
43	74.5	12.4	1248	3 US-08-882-046-6	Sequence 6, Appli
44	74.5	12.4	1248	4 US-09-566-047-6	Sequence 6, Appli
45	74	12.3	1068	1 US-08-537-210A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5765883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

;USES AND METHODS OF PRODUCING SAME

; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO:1:
; LENGTH: 2327
5455158-1

Query Match 100.0%; Score 600; DB 6; Length 2327;
Best Local Similarity 100.0%; Pred. No. 8.7e-55;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 4 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGFNCSKP 63
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
Db 64 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 109

RESULT 5

US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plummer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-09-016-366A-12

Query Match 100.0%; Score 600; DB 2; Length 2386;
Best Local Similarity 100.0%; Pred. No. 9e-55;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 32 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGFNCSKP 91
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 137

RESULT 6

US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246
; REFERENCE/DOCKET NUMBER: 92-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-551-356-2

Query Match 100.0%; Score 600; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 9.2e-55;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 32 QAQMVPQSPVAVSQSKPGCYDNGKHQYQINQWERTYLGNVLVCTCYGSGRGFNCSKP 91
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 106
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRICTI 137

RESULT 7

PCT-US93-12687-2
; Sequence 2, Application PC/TUS9312687
; GENERAL INFORMATION:
; APPLICANT: Iranli, Meher H.
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12687
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/998,271
; FILING DATE: 31-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 92-26PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-12687-2
Query Match 100.0%; Score 600; DB 5; Length 2446;
Best Local Similarity 100.0%; Pred. No. 9,2e-55;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QAOQMVQPSVAVSQKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 32 QAOQMVQPSVAVSQKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTI 106
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTI 137
RESULT 8
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOUT
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: WINTERHAGER, ELKE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-961-403-1
Query Match 99.3%; Score 596; DB 4; Length 2386;
Best Local Similarity 99.1%; Pred. No. 2.4e-54;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 QAOQMVQPSVAVSQKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 32 QAOQMVQPSVAVSQKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 91
Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTI 106
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGRI SCTI 137
RESULT 9
US-08-142-449B-14
; Sequence 14, Application US/08142449B
; Patent No. 5688104
; GENERAL INFORMATION:
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kawano, Genji
; APPLICANT: Sudo, Tetsuo
; APPLICANT: Kojima, Katsuaki
; TITLE OF INVENTION: Physiologically Active Protein and
; REFERENCE/DOCKET NUMBER: Hematopoietic Stem Cell Growth Agent
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nels T. Lippert, White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,449B
; FILING DATE: 24-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lippert, Nels T.
; REGISTRATION NUMBER: 25,888
; REFERENCE/DOCKET NUMBER: 1145358-304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 354-8113
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-142-449B-14
Query Match 27.1%; Score 162.5; DB 1; Length 188;
Best Local Similarity 39.0%; Pred. No. 1.2e-09;
Matches 32; Conservative 15; Mismatches 28; Indels 7; Gaps 4;
Qy 21 CYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRG-FNCESKPAEETCFDKYTGNTYRVG 78
Db 53 CHDNGVNYKIGEKWDQNGQNMWCTCLNGNGKEFKCDPH---EATCYD--DGKTVHG 107
Qy 79 DTYERPKDSMIWDCTCIGAGRG 100
Db 108 EQWQKEYLGAICSTCTCFGGORG 129

```

; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,854
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/Jp95/02425
; FILING DATE: 29-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 317721/1994
; FILING DATE: 29-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: HASHINO=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-836-854-9

Query Match 27.1%; Score 162.5; DB 2;
Best Local Similarity 39.0%; Pred. No. 3.5e-09;
Matches 32; Conservative 15; Mismatches 28;

QY 21 CYDNGKHGYOINQWERT-YLGNVLVCTCYGSGRG-FNCSES
DB 339 CHDNGVNYKIGKWRQSGNGQMMSCTCLGNKGKGFKCDPI
QY 79 DTYERKDSMIWDCCTICAGRG 100
DB 394 EQWQKEYLGAICSTCFGGQRG 415

RESULT 12
US-08-340-812-7
; Sequence 7, Application US/08340812
; Patent No. 5629291
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND
; TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRA
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO.
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,812

```

```
;
; FILING DATE: 17-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021,626
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BELLAS, CHRISTINE M.
; REGISTRATION NUMBER: 34,122
; REFERENCE/DOCKET NUMBER: P-LA 9442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-340-812-7

Query Match 16.7%; Score 100; DB 1; Length 42;
Best Local Similarity 40.5%; Pred. No. 0.00091;
Matches 15; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 22 YDNGKHYQINQOWERTYLGVLVCTCYGSGRGFNCS 58
||:||||:||||:||||:||||:||||:||||:
Db 4 YDDGKTYHVGEOQWKEYLGAISSSTSGGQGRWRSN 40

RESULT 13
US-08-459-064B-7
; Sequence 7, Application US/08459064B
; Patent No. 5747452
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: A METHOD OF MODULATING TUMOR CELL MIGRATION
; TITLE OF INVENTION: USING FIBRONECTIN TYPE III PEPTIDES
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 01-JUN-1995
; APPLICATION NUMBER: US/08/459,064B
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,462
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1543
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-421A-7

Query Match 16.7%; Score 100; DB 1; Length 42;
Best Local Similarity 40.5%; Pred. No. 0.00091;
Matches 15; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 22 YDNGKHYQINQOWERTYLGVLVCTCYGSGRGFNCS 58
||:||||:||||:||||:||||:||||:||||:
Db 4 YDDGKTYHVGEOQWKEYLGAISSSTSGGQGRWRSN 40

RESULT 14
US-08-460-421A-7
; Sequence 7, Application US/08460421A
; Patent No. 5837813
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
; TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,421A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,462
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; APPLICATION NUMBER: US 08/340,812
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1542
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-421A-7

Query Match 16.7%; Score 100; DB 2; Length 42;
Best Local Similarity 40.5%; Pred. No. 0.00091;
Matches 15; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 22 YDNGKHYQINQOWERTYLGVLVCTCYGSGRGFNCS 58
||:||||:||||:||||:||||:||||:||||:
Db 4 YDDGKTYHVGEOQWKEYLGAISSSTSGGQGRWRSN 40
```

```
;
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-459-064B-7

Query Match 16.7%; Score 100; DB 1; Length 42;
Best Local Similarity 40.5%; Pred. No. 0.00091;
Matches 15; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 22 YDNGKHYQINQOWERTYLGVLVCTCYGSGRGFNCS 58
||:||||:||||:||||:||||:||||:||||:
Db 4 YDDGKTYHVGEOQWKEYLGAISSSTSGGQGRWRSN 40

RESULT 14
US-08-460-421A-7
; Sequence 7, Application US/08460421A
; Patent No. 5837813
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: MORLA, ALEX
; TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
; TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES LLP
; STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,421A
; FILING DATE: 01-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/829,462
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/021,626
; FILING DATE: 16-FEB-1993
; APPLICATION NUMBER: US 08/340,812
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, CATHRYN A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1542
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-460-421A-7

Query Match 16.7%; Score 100; DB 2; Length 42;
Best Local Similarity 40.5%; Pred. No. 0.00091;
Matches 15; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 22 YDNGKHYQINQOWERTYLGVLVCTCYGSGRGFNCS 58
||:||||:||||:||||:||||:||||:||||:
Db 4 YDDGKTYHVGEOQWKEYLGAISSSTSGGQGRWRSN 40
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:54:46 ; Search time 38.8425 Seconds
(without alignments)
884.776 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQMVPQSPVAVSQSPKPG.....SMIWDCTCIGAGRGISCTI 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	100.0	259	10 US-09-940-235-4	Sequence 4, Appli
2	600	100.0	2328	14 US-10-171-311-64	Sequence 64, Appl
3	600	100.0	2328	14 US-10-236-031B-70	Sequence 70, Appl
4	600	100.0	2328	14 US-10-374-979-98	Sequence 98, Appl
5	600	100.0	2328	15 US-10-182-936A-98	Sequence 98, Appl
6	596	99.3	231	9 US-09-925-302-548	Sequence 548, App
7	596	99.3	231	10 US-09-925-302-548	Sequence 548, App
8	596	99.3	231	14 US-10-144-194A-52	Sequence 52, Appl
9	596	99.3	463	16 US-10-741-601-354	Sequence 354, App
10	596	99.3	657	16 US-10-741-601-359	Sequence 359, App
11	596	99.3	984	16 US-10-741-601-356	Sequence 356, App
12	596	99.3	2220	15 US-10-236-392-4	Sequence 4, Appli
13	596	99.3	2296	16 US-10-741-601-363	Sequence 363, App

14	596	99.3	2355	14	US-10-144-194A-104	Sequence 104, App
15	596	99.3	2355	15	US-10-360-101-235	Sequence 235, App
16	596	99.3	2355	15	US-10-447-161-3	Sequence 3, Appli
17	596	99.3	2355	16	US-10-734-564-94	Sequence 94, Appl
18	596	99.3	2355	16	US-10-741-601-357	Sequence 357, App
19	596	99.3	2355	16	US-10-741-601-366	Sequence 366, App
20	596	99.3	2386	10	US-09-961-403-1	Sequence 1, Appli
21	596	99.3	2386	16	US-10-741-601-360	Sequence 360, App
22	596	99.3	2386	17	US-10-788-792-206	Sequence 206, App
23	581	96.8	2320	15	US-10-236-392-2	Sequence 2, Appli
24	577	96.2	2320	14	US-10-279-733-8	Sequence 8, Appli
25	248	41.3	164	9	US-09-925-301-1270	Sequence 1270, Ap
26	167.5	27.9	847	16	US-10-741-601-361	Sequence 361, App
27	164.5	27.4	1286	16	US-10-741-601-362	Sequence 362, App
28	162.5	27.1	163	16	US-10-734-564-115	Sequence 115, App
29	162.5	27.1	163	17	US-10-770-668-38	Sequence 38, Appl
30	162.5	27.1	693	16	US-10-741-601-364	Sequence 364, App
31	162.5	27.1	793	14	US-10-171-311-62	Sequence 62, Appl
32	162.5	27.1	1259	16	US-10-741-601-365	Sequence 365, App
33	162.5	27.1	1315	16	US-10-741-601-358	Sequence 358, App
34	162.5	27.1	1341	16	US-10-741-601-355	Sequence 355, App
35	162.5	27.1	1348	16	US-10-741-601-353	Sequence 353, App
36	161.5	26.9	170	15	US-10-264-049-3220	Sequence 3220, Ap
37	148.5	24.8	86	15	US-10-424-599-230915	Sequence 230915,
38	143	23.8	343	9	US-09-934-706-1	Sequence 1, Appli
39	143	23.8	400	9	US-09-934-706-5	Sequence 5, Appli
40	143	23.8	501	9	US-09-934-706-4	Sequence 4, Appli
41	93	15.5	895	15	US-10-294-006-35	Sequence 35, Appl
42	90	15.0	228	15	US-10-344-634-4	Sequence 4, Appli
43	90	15.0	285	15	US-10-344-634-16	Sequence 16, Appl
44	90	15.0	386	15	US-10-344-634-14	Sequence 14, Appl
45	85	14.2	1568	15	US-10-263-929-117	Sequence 117, App

ALIGNMENTS

RESULT 1

US-09-940-235-4
; Sequence 4, Application US/09940235
; Publication No. US2003005921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chaït
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4

Query Match 100.0%; Score 600; DB 10; Length 259;

Best Local Similarity 100.0%; Pred. No. 5.6e-58;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAQMVPQSPVAVSQSPKPGCYDNGKHQYQINQWERTYLGVLVCTCYGSGRGNCBCKP 60


```
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 98
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-936A-98

Query Match          100.0%; Score 600; DB 15; Length 2328;
Best Local Similarity 100.0%; Pred. No. 6.9e-57;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
Db      5  QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 64

Qy      61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCTCIGAGRGRICTI 106
Db      65 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCTCIGAGRGRICTI 110

RESULT 6
US-09-925-302-548
; Sequence 548, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-548

Query Match          99.3%; Score 596; DB 9; Length 231;
Best Local Similarity 99.1%; Pred. No. 1.4e-57;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
Db      61 QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 120

Qy      61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCTCIGAGRGRICTI 106
Db      121 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCTCIGAGRGRICTI 166

RESULT 7
US-09-925-302-548
; Sequence 548, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 548
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (205)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (212)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-548

Query Match          99.3%; Score 596; DB 10; Length 231;
Best Local Similarity 99.1%; Pred. No. 1.4e-57;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 60
Db      61 QAOQMVQPSVAVSQSPGKCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCSKP 120

Qy      61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCTCIGAGRGRICTI 106
Db      121 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCTCIGAGRGRICTI 166

RESULT 8
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-194A-52

Query Match          99.3%; Score 596; DB 14; Length 463;
Best Local Similarity 99.1%; Pred. No. 3e-57;
```

```

Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 QAQMWPQSPVAVSQSPKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCESKP 60
    |||
Db 68 QAQMWPQSPVAVSQSPKPGCYDNGKHQYQINQOWERTYLGVLVCTCYGSGRGFNCESKP 127
    |||
QY 61 EABETCFDKTGTNTYRVGDTYVERPKDSMIWMDCTCIGAGRGRISCTI 106
    |||
Db 128 EABETCFDKTGTNTYRVGDTYVERPKDSMIWMDCTCIGAGRGRISCTI 173
    |||

RESULT 9
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

```

```

RESULT 11
US-10-741-601-356
; Sequence 356, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 28415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-356

Query Match          99.3%; Score 596; DB 16; Length 984;
Best Local Similarity 99.1%; Pred. No. 7.2e-57;
Matches 105; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  QAOQMVPQSPVAVSQSKPGCYDNGKHYYQINQWERTYLGNLVLCCTCYGSGRGFNCE
Db      32  QAOQMVPQSPVAVSQSKPGCYDNGKHYYQINQWERTYLGNLVLCCTCYGSGRGFNCE

Qy      61  EASETCFDKYTGNTYRVGDTYERPDKDSMTIWDCTCTGAGRGRIISCTI 106
Db      92  EASETCFDKYTGNTYRVGDTYERPDKDSMTIWDCTCTGAGRGRIISCTI 137

```

Search completed: November 4, 2004, 00:15:28
Job time : 39.8425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:44:40 ; Search time 155.188 Seconds
(without alignments)
756.808 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQMVPQSPVAVSQSKPG.....SMWDTCTGAGRGRICTI 106

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
- 32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
- 33: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
- 34: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
- 35: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
- 36: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB	ID	Description
1	600	100.0	259	18	US-09-471-349-4	Sequence 4, Appli
2	600	100.0	259	24	US-09-940-235-4	Sequence 4, Appli
3	600	100.0	259	32	US-10-631-558-4	Sequence 4, Appli
4	600	100.0	657	19	US-09-581-651B-36	Sequence 36, Appl
5	600	100.0	657	19	US-09-581-651C-36	Sequence 36, Appl
6	600	100.0	675	19	US-09-581-651B-1	Sequence 1, Appli
7	600	100.0	675	19	US-09-581-651C-1	Sequence 1, Appli
8	600	100.0	719	19	US-09-581-651B-37	Sequence 37, Appl
9	600	100.0	719	19	US-09-581-651C-37	Sequence 37, Appl
10	600	100.0	720	19	US-09-581-651B-44	Sequence 44, Appl
11	600	100.0	720	19	US-09-581-651C-44	Sequence 44, Appl
12	600	100.0	2324	1	PCT-US00-12993-3	Sequence 3, Appli
13	600	100.0	2324	1	PCT-US00-12993-3	Sequence 3, Appli
14	600	100.0	2324	3	US-07-714-134-2	Sequence 2, Appli
15	600	100.0	2324	3	US-07-714-134B-2	Sequence 2, Appli
16	600	100.0	2324	28	US-10-276-002-3	Sequence 3, Appli
17	600	100.0	2324	28	US-10-276-006-3	Sequence 3, Appli
18	600	100.0	2327	18	US-09-492-971A-15	Sequence 15, Appl
19	600	100.0	2328	1	PCT-US02-10824-196	Sequence 196, App
20	600	100.0	2328	1	PCT-US02-14877A-677	Sequence 677, App
21	600	100.0	2328	1	PCT-US02-15982-98	Sequence 98, Appl
22	600	100.0	2328	1	PCT-US02-15982A-98	Sequence 98, Appl
23	600	100.0	2328	1	PCT-US02-18638A-64	Sequence 64, Appl
24	600	100.0	2328	27	US-10-171-311-64	Sequence 64, Appl
25	600	100.0	2328	27	US-10-182-936A-98	Sequence 98, Appl
26	600	100.0	2328	28	US-10-236-031B-70	Sequence 70, Appl
27	600	100.0	2328	28	US-10-240-851-98	Sequence 98, Appl
28	600	100.0	2328	29	US-10-374-979-98	Sequence 98, Appl
29	600	100.0	2328	30	US-10-477-173-677	Sequence 677, App
30	600	100.0	2328	30	US-10-477-238A-677	Sequence 677, App
31	600	100.0	2328	32	US-10-680-287A-677	Sequence 677, App
32	600	100.0	2328	36	US-60-290-071-98	Sequence 98, Appl
33	600	100.0	2359	1	PCT-US01-14827-12148	Sequence 12148, A
34	600	100.0	2386	14	US-09-016-366-12	Sequence 12, Appl
35	600	100.0	2446	1	PCT-US01-14827-12147	Sequence 12147, A
36	600	100.0	2474	1	PCT-US01-08631-52638	Sequence 52638, A
37	600	100.0	2477	18	US-09-446-274A-7	Sequence 7, Appli
38	600	100.0	2477	18	US-09-446-274B-7	Sequence 7, Appli
39	596	99.3	231	1	PCT-US00-05918-548	Sequence 548, App
40	596	99.3	231	24	US-09-925-302-548	Sequence 548, App
41	596	99.3	463	1	PCT-US02-31287-52	Sequence 52, Appl
42	596	99.3	463	27	US-10-144-194A-52	Sequence 52, Appl
43	596	99.3	463	30	US-10-491-566-52	Sequence 52, Appl
44	596	99.3	642	1	PCT-US03-40977-354	Sequence 354, App
45	596	99.3	642	1	PCT-US03-40978-1066	Sequence 1066, Ap

ALIGNMENTS

RESULT 1
US-09-471-349-4
; Sequence 4, Application US/09471349
; GENERAL INFORMATION:
; APPLICANT: Sahni, Girish
; APPLICANT: Kumar, Rajesh
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
; PREPARATION OF SAID PROTEINS
; FILE REFERENCE: 07064/009001
; CURRENT APPLICATION NUMBER: US/09/471,349
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 24

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-471-349-4

Query Match          100.0%; Score 600; DB 18; Length 259;
Best Local Similarity 100.0%; Pred. No. 5.4e-57;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

RESULT 2
US-09-940-235-4
; Sequence 4, Application US/09940235
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4

Query Match          100.0%; Score 600; DB 24; Length 259;
Best Local Similarity 100.0%; Pred. No. 5.4e-57;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

RESULT 3
US-10-631-558-4
; Sequence 4, Application US/10631558
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
```

```
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-558-4

Query Match          100.0%; Score 600; DB 32; Length 259;
Best Local Similarity 100.0%; Pred. No. 5.4e-57;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106

RESULT 4
US-09-581-651B-36
; Sequence 36, Application US/09581651B
; GENERAL INFORMATION:
; APPLICANT: University of Dundee, University of Dundee
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses Thereof
; FILE REFERENCE: 350013-72
; CURRENT APPLICATION NUMBER: US/09/581,651B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 36
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651B-36

Query Match          100.0%; Score 600; DB 19; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.5e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 32 QAAQMVQPSPVAVSQSGKPGCYDNGKHQYQINQOWERTYLGNVLVCTCYGSGRGFNCSKP 91

Qy 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 106
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRGISCTI 137

RESULT 5
US-09-581-651C-36
; Sequence 36, Application US/09581651C
; GENERAL INFORMATION:
; APPLICANT: Schor, Seth Lawrence
; APPLICANT: Schor, Ana Maria
```

; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses
; FILE REFERENCE: ERPO1.003APC
; CURRENT APPLICATION NUMBER: US/09/581,651C
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: GB 9726539.1
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651C-36

Query Match 100.0%; Score 600; DB 19; Length 657;
Best Local Similarity 100.0%; Pred. No. 1.5e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 60
Db 32 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 91

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
Db 92 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 137

RESULT 6
US-09-581-651B-1
; Sequence 1, Application US/09581651B
; GENERAL INFORMATION:
; APPLICANT: University of Dundee, University of Dundee
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses Thereof
; FILE REFERENCE: 350013-72
; CURRENT APPLICATION NUMBER: US/09/581,651B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651B-1

Query Match 100.0%; Score 600; DB 19; Length 675;
Best Local Similarity 100.0%; Pred. No. 1.6e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 60
Db 50 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 109
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 155

RESULT 7
US-09-581-651C-1
; Sequence 1, Application US/09581651C
; GENERAL INFORMATION:
; APPLICANT: Schor, Seth Lawrence
; APPLICANT: Schor, Ana Maria
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses
; FILE REFERENCE: ERPO1.003APC
; CURRENT APPLICATION NUMBER: US/09/581,651C
; CURRENT FILING DATE: 2000-10-10

; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: GB 9726539.1
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651C-1

Query Match 100.0%; Score 600; DB 19; Length 675;
Best Local Similarity 100.0%; Pred. No. 1.6e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 60
Db 50 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 109

QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 155

RESULT 8
US-09-581-651B-37
; Sequence 37, Application US/09581651B
; GENERAL INFORMATION:
; APPLICANT: University of Dundee, University of Dundee
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses Thereof
; FILE REFERENCE: 350013-72
; CURRENT APPLICATION NUMBER: US/09/581,651B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (676)..(676)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (679)..(679)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (683)..(683)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (717)..(717)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-581-651B-37

Query Match 100.0%; Score 600; DB 19; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.7e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 60
Db 50 QAAQMVQPSVAVSQSKPCYDNGKHQYQINQWERTYLGNNLVCTCYGSGRGFNCESKP 109
QY 61 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYERPKDSMIWDCTCIGAGRISCTI 155

```
RESULT 9
US-09-581-651C-37
; Sequence 37, Application US/09581651C
; GENERAL INFORMATION:
; APPLICANT: Schor, Seth Lawrence
; APPLICANT: Schor, Ana Maria
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: EP01.003APC
; CURRENT APPLICATION NUMBER: US/09/581,651C
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: GB 9726539.1
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 719
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 676, 679, 683, 717
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-581-651C-37

Query Match 100.0%; Score 600; DB 19; Length 719;
Best Local Similarity 100.0%; Pred. No. 1.7e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 60
Db 50 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 109

Qy 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 155

RESULT 10
US-09-581-651B-44
; Sequence 44, Application US/09581651B
; GENERAL INFORMATION:
; APPLICANT: University of Dundee, University of Dundee
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses Thereof
; FILE REFERENCE: 350013-72
; CURRENT APPLICATION NUMBER: US/09/581,651B
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 44
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651B-44

Query Match 100.0%; Score 600; DB 19; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.7e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 60
Db 50 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 109

Qy 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 155

RESULT 11
US-09-581-651C-44
; Sequence 44, Application US/09581651C
; GENERAL INFORMATION:
; APPLICANT: Schor, Seth Lawrence
; APPLICANT: Schor, Ana Maria
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: EP01.003APC
; CURRENT APPLICATION NUMBER: US/09/581,651C
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: GB 9726539.1
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651C-44

Query Match 100.0%; Score 600; DB 19; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.7e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 60
Db 50 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 109

Qy 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 155
```

```
US-09-581-651C-44
; Sequence 44, Application US/09581651C
; GENERAL INFORMATION:
; APPLICANT: Schor, Seth Lawrence
; APPLICANT: Schor, Ana Maria
; TITLE OF INVENTION: Polypeptides, Polynucleotides and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: EP01.003APC
; CURRENT APPLICATION NUMBER: US/09/581,651C
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: PCT/GB98/03766
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: GB 9726539.1
; PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-581-651C-44

Query Match 100.0%; Score 600; DB 19; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.7e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 60
Db 50 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 109

Qy 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 106
Db 110 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 155

RESULT 12
PCT-US00-12961-3
; Sequence 3, Application PC/TUS0012961
; GENERAL INFORMATION:
; APPLICANT: Williams, David A.
; APPLICANT: Bradford, Gillian B.
; APPLICANT: Dutt, Parmesh
; APPLICANT: Yoder, Mervin C.
; APPLICANT: Advanced Research and Technology Institute
; TITLE OF INVENTION: Methods for Enriching for Quiescent Cells in
; TITLE OF INVENTION: Hematopoietic Cell Populations
; FILE REFERENCE: IUI04
; CURRENT APPLICATION NUMBER: PCT/US00/12961
; CURRENT FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 3
; LENGTH: 2324
; TYPE: PRT
; ORGANISM: homo sapiens
PCT-US00-12961-3

Query Match 100.0%; Score 600; DB 1; Length 2324;
Best Local Similarity 100.0%; Pred. No. 6.4e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 60
Db 1 QAAQWVQPSVAVSQSGKPCYDNGKHQYQINQWERTYLGNIWLVCTCYGSGRGFNCSKP 60

Qy 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYRPPKDSMIWDCCTCIGAGRGISCTI 106

RESULT 13
PCT-US00-12993-3
; Sequence 3, Application PC/TUS0012993
```



```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-714-134-2

Query Match 100.0%; Score 600; DB 3; Length 2324;
Best Local Similarity 100.0%; Pred. No. 6.4e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSPKPCYDNGKHVQIQNQWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 1 QAOQWVQPSVAVSQSPKPCYDNGKHVQIQNQWERTYLGNVLVCTCYGSGRGFNCSKP 60

QY 61 EAEETCFDKYTGNTYRVGDTYVERPKDSMIWDCTCIGAGRGRISCTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYVERPKDSMIWDCTCIGAGRGRISCTI 106

RESULT 15
US-07-714-134B-2
; Sequence 2, Application US/07714134B
; GENERAL INFORMATION:
; APPLICANT: GOLD, LESLIE I.
; APPLICANT: ROSTAGNO, AGUEDA A.
; APPLICANT: BARON, MARTIN
; APPLICANT: CAMPBELL, IAIN D.
; APPLICANT: WILLIAMS, MICHAEL J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy & Neimark
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/714,134B
; FILING DATE: 19910614
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Guy K.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: GOLD 1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-714-134B-2

Query Match 100.0%; Score 600; DB 3; Length 2324;
Best Local Similarity 100.0%; Pred. No. 6.4e-56;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QAOQWVQPSVAVSQSPKPCYDNGKHVQIQNQWERTYLGNVLVCTCYGSGRGFNCSKP 60
Db 1 QAOQWVQPSVAVSQSPKPCYDNGKHVQIQNQWERTYLGNVLVCTCYGSGRGFNCSKP 60

QY 61 EAEETCFDKYTGNTYRVGDTYVERPKDSMIWDCTCIGAGRGRISCTI 106
Db 61 EAEETCFDKYTGNTYRVGDTYVERPKDSMIWDCTCIGAGRGRISCTI 106

```

Search completed: November 4, 2004, 00:11:21
Job time : 156.188 secs

```
US-09-940-235-4_COPY_1_106 (1-106) x AR274901 (1-7679)

QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTAGTCAAAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db 134 AATGTGTTGGTTGTACTGTTATGAGGAGCCGAGGTTTAACTCGCAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysTleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 2
LOCUS AR380744 7680 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1289 from patent US 6607879.
ACCESSION AR380744
VERSION AR380744.1 GI:40088378
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7680)
AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
TITLE Compositions for the detection of blood cell and immunological
JOURNAL response gene expression
PATENT: US 6607879-A 1289 19-AUG-2003;
FEATURES
source
location/Qualifiers
1..7680
/mol_type="genomic DNA"

ORIGIN
Alignment Scores: 1.16e-59 Length: 7680
Pred. No.: 600.00 Matches: 106
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-940-235-4_COPY_1_106 (1-106) x AR380744 (1-7680)

QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTAGTCAAAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db 134 AATGTGTTGGTTGTACTGTTATGAGGAGCCGAGGTTTAACTCGCAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
```

```
QY 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysTleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 3
LOCUS AX277596 7680 bp DNA linear PAT 01-NOV-2001
DEFINITION Sequence 75 from Patent WO0177327.
ACCESSION AX277596
VERSION AX277596.1 GI:16604795
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Carulli,J.P., Little,R.D., Recker,R.R. and Johnson,M.L.
TITLE The high bone mass gene of Ilq13.3
JOURNAL Patent: WO 0177327-A 75 18-OCT-2001;
Genome Therapeutics Corporation (US)
FEATURES
source
location/Qualifiers
1..7680
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 1.16e-59 Length: 7680
Pred. No.: 600.00 Matches: 106
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-940-235-4_COPY_1_106 (1-106) x AX277596 (1-7680)

QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGGTGGCTGTAGTCAAAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db 134 AATGTGTTGGTTGTACTGTTATGAGGAGCCGAGGTTTAACTCGCAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysTleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 4
LOCUS AX335368 7680 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 5877 from Patent WO0194629.
ACCESSION AX335368
VERSION AX335368.1 GI:18126087
KEYWORDS
```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 5877 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
source
1. 7680
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.16e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x AX335368 (1-7680)
Qy 1_GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGGTTCAGCCCAAGTCCCGCGTGTCTAGTCAAAAGCAAGCCGGT 73
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTropGluArgThrTyrLeuGly 40
Db 74 TGTATTGACATGGAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
Qy 41 AnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db 134 AATGTTGTTTGTACTTGTATGGAGGAGCGAGGTTTAACTGCGAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTGCACAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
Qy 81 TyrGluArgProIysAspSerMetIleTropAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCCAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331
RESULT 5
HSFIB1
LOCUS HSFIB1 7680 bp mRNA linear PRI 01-OCT-1999
DEFINITION Human mRNA for fibronectin (FN precursor).
ACCESSION X02761 K00055 K00799 K02273 X00307 X00739
VERSION X02761.1 GI:31396
KEYWORDS alternate splicing; fibronectin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kornblitt, A.R., Vibe-Pedersen, K. and Baralle, F.E.
TITLE Isolation and characterization of cDNA clones for human and bovine
fibronectins
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 80 (11), 3218-3222 (1983)
MEDLINE 83221567
PUBMED 6304699
REFERENCE 2
AUTHORS Kornblitt, A.R., Vibe-Pedersen, K. and Baralle, F.E.
TITLE Human fibronectin: molecular cloning evidence for two mRNA species

differing by an internal segment coding for a structural domain
EMBO J. 3 (1), 221-226 (1984)
84158533
6200322
REFERENCE 3
AUTHORS Kornblitt, A.R., Vibe-Pedersen, K. and Baralle, F.E.
TITLE Human fibronectin: cell specific alternative mRNA splicing
generates polypeptide chains differing in the number of internal
repeats
Nucleic Acids Res. 12 (14), 5853-5868 (1984)
84272258
6462919
REFERENCE 4 (bases 1 to 7680)
AUTHORS Kornblitt, A.R., Umezawa, K., Vibe-Pedersen, K. and Baralle, F.E.
TITLE Primary structure of human fibronectin: differential splicing may
generate at least 10 polypeptides from a single gene
EMBO J. 4 (7), 1755-1759 (1985)
85284965
2992939
COMMENT On or before Jan 13, 2004 this sequence version replaced gi:182680,
gi:182681, gi:182698.
FEATURES
source
Location/Qualifiers
1. 7680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
<1..6988
/codon_start=2
/product="fibronectin precursor"
/protein_id="CAA26536.1"
/db_xref="GI:31397"
/db_xref="GOA:P02751"
/translation="KSKRQAOQMVQPSVAVSQSPKPCYDNGKHQVINOQWERTYILG
NVLCTCYGSGRGNCFNCEKPEABETCFDKTIGNTYRVDGTYRKPQKQIMWDCICIGAG
RGRISCIANRCHEGQSYKIGDTWRPHETGGYLMCEVCVLCNGKGEWTKCPAEEKCF
DHAAGTSYVVGTEWPKYQGMVMVDCICLGEHSGYRITCTSRNRCNDQDTRTSYRIGDT
WSKCHNRGNLLOCICTGNRGKWKCRHTSVQTSQSGGPTDVAATVQPOPHQPP
PYGCHVDSGVYVSGMOLKTOGNKMLCTCLGNGVSCQETAVTQYGNLNGEPCV
LPTYNGRFYSCCTEGRDGHLWCSTTSNYEQDQKYSFCTDHTVLVQDQGNNGAL
CHFFPLNYHTDCTSEGRDNKMGCTTQNDQKFGCPMAHEEICTNTEGVM
YRTGDQKHQNDHMRCTCVNGRGWETCIAYSLRQDCIVDDIYNNVDYTFHKKH
EGHMLNCTCFQGRGWKCDPVDQCDSETGYQIGDSWEKYHGVRYCYCYGRG
IGWHQCPLOTYPSSGPVEVITETPSQNSHPIONWAPQPSHKYILRWPKNSV
GRWEATIPHLNSYIKGLKGVVYEGOLISIQXQHOEVTEFDTTSTSTPTVNS
TVGTETPSPVATSEVTEITASSFVSVWSASDTVSGFRVEYELSEEGDPOYLD
LPSTATSVNIPDLLGRKYIVNVYQISEDEGSQLILSTQTTAPDAPPDPTVDQDDT
SIVVWSRFPQAPITGYRIYVSPSGSSTELNLPETANSVTLSDLPQGVQNYITVAV
EENQESTPVVIOQETTPRSDTVPSGPRDLQFVEVDVKVTIMWTPESAGVITGYVAV
IPVNLPEHGRLPISRNTPAEVTGLSPGVYVFKVAVSHGSKPLTAQOQTKLDA
PTNLOFNETDSTVLYRWTPPRAQITGYRLTVGLTRRGQPRQVYNGSPSVKYLRLNQ
PASEYTVSLVAKNGESPATGVFTLQPGSSIPPINTEVETITVITWIPAPRIGF
KLGRPSQGEAPREVTSQSGSIWSGLTPGVYVYITVQLRDQERDAPVKNVYTP
LSPPTNLHLEANPDTGLTVSWERSTPDITGYRIITTTNGQGNLSLEEVHADQSS
CTFDNLSPLGVNVSVTVKDKESVPSIDTIIPAVPPDLDRTNIGDPMVMTWAP
PSSIDITNLVRYSPVKNEDEVAELSISSPDNAVLTNLLPGTEYVYVSVSVYEQHES
TPLRQKQGLDPTGIDFSDITANSFTVHIAPIRATITGYRIHHPHSPGSPREDR
VPHRSNITLITNLTPTGYEVSIVALNGREESPLLQSQSTSDVSDPRDLEVAATPTS
LLISDPAVATVYRITVYGTGCGSPQVEFTVPSGKSTATISGLKPGVDYITVYAV
NGPGPTKTATAGDQETMEIDKPSQMVTDQDNSISVKWLPSSSPGVGTGYVTTTPK
LAFTDVEDSIKIAMESPOGVSEYVTVSSPEDGHELFAPDGEDAELOGLRPG
SEYTVSVLALHDMESQPLIGTOSTAIPTADLTKQTQVTPSLAQWTPNVQLQTYR
VRVTPKEKTPMKBEINLADSSVSVGLMVAATKYVSVVALKDTLSTSPAQGVTTTL
ENYPPRRARVTDATETITISWRTKETITGQVDVAFANGQTPORTIKPDVRSYT
ITGLQPTDYKIYLTINDNARSPPVIDASTADPSNLRLATPLNSLNSVQPPR
ARTIGYIIVKEKPGSPREVPRPGVTEATITGLEPGTEYTYVIALKNKQKSEPL
IGRKTDELPLQVLTPHNLHPEILDVSTVOKTFVTHPGVDYTGNGILQPTGSSQO
PSVQOMIIEFHGFRRTPTTATPIRHRPPTPPYPPNVQEALESQTTISWAPQDTSY
IISCHPVDDEEPLQFVPTSATLGLTRGATYNIIVKALQDQKHKRVSEVTV
GNSNEGLNQPTDSDCFDPTVSHYAVGDEWERMESGFKLCCQLGFGSGHRCFDS
RWCHDNGVYKIGEKWDQNGQMSCTCLNGKGEFKCDPHEATCYDDGKTYHVG
QWQKEYLGAICSCTCFGGQGRWRCNCRPGGEPSPGTTGQSYNQSYQRYHQTWTN

mat_peptide
14. .6985
/product="fibronectin precursor"
/note="(pot. aa 1-2324)"
74. .736
/note="region of internal homology I (5 subunits); domain with fibrin-, heparin-, s.aureus-binding activity"
842. .952
/note="region of internal homology I (1 subunit); domain with collagen-binding activity"
953. .1327
/note="region of internal homology II (2 subunits) domain with collagen-binding activity"
1328. .1744
/note="region of internal homology I (3 subunits) domain with collagen-binding activity"
1745. .2020
/note="region of internal homology III (1 subunit)"
2075. .5893
/note="region of internal homology III (14 subunits)"
2347. .3175
/note="domain with DNA-binding activity"
4490. .4501
/note="cell binding site"
4811. .5080
/note="ED region for alternative splicing"
5081. .5893
/note="domain with heparin-binding activity"
5894. .6160
/note="III CS region for alternative splicing"
6161. .6373
/note="region of internal homology III (1 subunit)"
6442. .6838
/note="region of internal homology I (3 subunits); domain with fibrin-binding activity"
7679

ORIGIN

Alignment Scores:
Pred. No.: 1.16e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x HSF1B1 (1-7680)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGGTTCAGCCCCAGTCCCGTGGCTGTTCAGTCAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTGTTGGTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACCTGGGAACACTTACCAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCTCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGAGGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 6

Al14133
LOCUS Fibronectin gene.
ACCESSION Al14133
VERSION Al14133.1 GI:490127
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7705)
AUTHORS Baralle, F.E.
TITLE Fibronectins
JOURNAL Patent: EP 0207751-A 2 07-JAN-1987;
Delta Biotechnology Limited
FEATURES
source
1. .7705
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature 5..6985
/note="CDS, fibronectin gene"
misc_feature 647..649
/note="the sequence given is TGA, but the codon is shown in the specification as encoding Glycine for which the correct codon is GGN"

ORIGIN

Alignment Scores:
Pred. No.: 1.16e-59 Length: 7705
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x Al14133 (1-7705)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGGTTCAGCCCCAGTCCCGTGGCTGTTCAGTCAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTGTTGGTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACCTGGGAACACTTACCAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCTCCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGAGGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331
RESULT 7
AR034630
LOCUS Sequence 16 from patent US 5869616.
ACCESSION AR034630
VERSION AR034630.1 GI:5950235
KEYWORDS Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7705)

ORIGIN

Alignment Scores:		
Pred. No.:	1.16e-59	Length: 7705
Score:	600.00	Matches: 106
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	6	Gaps: 0

US-09-940-235-4 COPY 1 106 (1-106) X I70110 (1-7705)

Qy		1	GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	20
Dd		14	CAGGCTCAGCAAAATGGTTTCAGCCCGGTCCCGGTGGCTGTCAAGTCAAAGCAAGCCCGGT	73
Qy		21	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrIeuGly	40
Dd		74	TGTTATGACAAATGGAAAAACATTATCAGATAAATCAACAGTGGGAGCGGACCCTACCTAGGT	133
Qy		41	AsnValIeuValCysThrCystTyrGlyGlySerArgGlyPheAsnCysGlnSerLysPro	60
Dd		134	AATGTGTGGTTGTGTACTTGTTATGGAGAAGCCGAGGTTTTTAATCGCGAAAGTAAACCT	193
Qy		61	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	80
Dd		194	GAAGCTGAAGAGACTTGCTTTGCACAAGTACACTGGGAACACTTACCGAGTGGGTGACACT	253
Qy		81	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly	100
Dd		254	TATGAGCGTCTCAAAGACTCCCATGATCTGGGACTGTACTGCATCGGGGCTGGCGGAGG	313
Qy		101	ArgIleSerCysThrIle	106
Dd		314	AGAAATAAGCTGTACCATC	331

RESULT 10

AR364992	AR364992	7705 bp	DNA	linear	PAT 03-SEP-2003
LOCUS	Sequence 2	from patent US 5455158.			
DEFINITION	Sequence				
ACCESSION	AR364992				
VERSION	AR364992.1	GI:34428226			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
	Unclassified.				
REFERENCE	1 (bases 1 to 7705)				
AUTHORS	Vogel,T., Levanon,A., Werber,M.M., Guy,R. and Panet,A.				
TITLE	Fibrin binding domain polypeptides and uses and methods of producing same				
JOURNAL	Patent: US 5455158-A 2 03-OCT-1995;				

ORIGIN

Alignment Scores:		
Pred. No.:	1.16e-59	Length: 7705
Score:	600.00	Matches: 106
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	6	Gaps: 0

U.S.-09-940-235-4 COPY 1 106 (1-106) x AR364992 (1-7705)

Qy 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSertysProGly 20
Db 14 CAGGCTCAGCAATGTTT CAGCCCCAGTCTCCCGGTGGCTGT CAGTCAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40

Db	74	TGTTATGACAAATGGAACACATATCAGATAAATCAACAGTGGAGCGGACCTTACCTAGGT	133
Qy	41	AsnValIeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerIysPro	60
Db	134	AATGTGTTGGTTTGTTACTGTGTTATCGAGAGCCGAGGTTTTAACTGCCGAAAGTAAACCT	193
Qy	61	GluAlaGluGluThrCysPheAspIysTyrThrGlyAsnThrTyrArgValGlyAspThr	80
Db	194	GAAGCTGAAGAGACTTGCTTTGACAAAGTACACTGGGAACACTTACCGATGGGTGACACT	253
Qy	81	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly	100
Db	254	TATGAGCGTCCTAAAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGCGGAGG	313
Qy	101	ArgIleSerCysThrIle	106
Db	314	AGAATAAGCTGTACCATC	331

RESULT 11

AR051657	7803 bp	DNA	linear	PAT 29-SEP-1999
LOCUS				
DEFINITION	Sequence 1 from patent	US 5830700.		
ACCESSION	AR051657			
VERSION	AR051657.1	GI:5975021		
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	Unclassified.			
AUTHORS	1 (bases 1 to 7803)			
TITLE	Irani, M.			
JOURNAL	Hybrid proteins having cross-linking and tissue-binding activities			
FEATURES	Patent: US 5830700-A 1 03-NOV-1998;			
	Location/Qualifiers			
source	1..7803			

ORIGIN

Alignment Scores:	
Pred. No.:	1.18e-59
Score:	600.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	7803
Matches:	106
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

U.S.-09-940-235-4 COPY 1 106 (1-106) x AP051657 (1-7803)

Qy	1	GlnAlaGlnMetValGlnProGlnSerProValAlaIleSerGlnSerLysProGly	20
Dd	99	CAGGCTCAGCAAAATGGTTTCAGCCCGTCCCAGTGCGTGTTCAGTCAAAGCAAGCCCGGT	158
Qy	21	CysTyrAspAsnGlyLysHisTyrGlnIleLeuGlnInTrpGluArgThrTyrLeuGly	40
Dd	159	TGTTATTGACAATGGAAAAACACTATCAGATAAATCAACACTGGGAGCGGACCCTACCTAGGT	218
Qy	41	AsnValIleuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGlnSerLysPro	60
Dd	219	AATGTGTGGTTGTGTACTTGTATGAGGAAGCCGAGGTTTTAACTGCCGAAAGATAAACCT	278
Qy	61	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	80
Dd	279	GAAGCTGNAGAGACTTGTGTTGACAAGTACACTTGGGAACACTTACGAGTGGGTGACACT	338
Qy	81	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArggly	100
Dd	339	TATCAGCGTCTCAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGCAGGG	398
Qy	101	ArgIleSerCysThrIle	106
Dd	399	AGAATAAGCTGTACCATC	416

RESULT 12

CQ731570
LOCUS CQ731570 706 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 17504 from Patent WO02068579.
ACCESSION CQ731570
VERSION CQ731570.1 GI:42308500
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 17504 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source 1..706
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2,65e-60 Length: 706
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x CQ731570 (1-706)
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 94 CAGGCTCAGCAATGGTTTCAGCCCACTCCCGGTGGCTGTCTAGTCAAGCAAGCCGGT 153
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 154 TGTATGACAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGGACTACCTAGGC 213
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 214 AATGCGTGGTTGTACTTGTATGAGGAGGAGCGAGGTTTAACTCGGAGAGTAACCT 273
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 274 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGCAACACTTACCGAGTGGGTGACACT 333
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 334 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGG 393
Qy 101 ArgIleSerCysThrIle 106
Db 394 AGAATAAGCTGTACCATC 411
RESULT 13
AX003229
LOCUS AX003229 2147 bp DNA linear PAT 24-AUG-2000
DEFINITION Sequence 2 from Patent WO991233.
ACCESSION AX003229
VERSION AX003229.1 GI:9927082
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Schor, S.L. and Schor, A.M.
TITLE Polypeptides, polynucleotides and uses thereof
JOURNAL Patent: WO 991233-A 2 24-JUN-1999;

SCHOR SETH LAWRENCE (GB); UNIV DUNDEE (GB)
FEATURES
source 1..2147
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 8,72e-60 Length: 2147
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x AX003229 (1-2147)
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 150 CAGGCTCAGCAATGGTTTCAGCCCACTCCCGGTGGCTGTCTAGTCAAGCAAGCCGGT 209
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 210 TGTATGACAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGGACTACCTAGGC 269
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 270 AATGCGTGGTTGTACTTGTATGAGGAGGAGCGAGGTTTAACTCGGAGAGTAACCT 329
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 330 GAAGCTGAAGAGACTTGTCTTTCAGCAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 389
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 390 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGG 449
Qy 101 ArgIleSerCysThrIle 106
Db 450 AGAATAAGCTGTACCATC 467
RESULT 14
BD137021
LOCUS BD137021 2147 bp DNA linear PAT 18-SEP-2002
DEFINITION Polypeptides, polynucleotides and uses thereof.
ACCESSION BD137021
VERSION BD137021.1 GI:23231966
KEYWORDS JP 2002508179-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Schor, S.L. and Schor, A.M.
TITLE Polypeptides, polynucleotides and uses thereof
JOURNAL Patent: JP 2002508179-A 1 19-MAR-2002;
UNIVERSITY OF DUNDEE
COMMENT OS Homo sapiens (human)
PN JP 2002508179-A/1
PD 19-MAR-2002
PF 15-DEC-1998 JP 2000539133
PR 16-DEC-1997 GB 9726539.1
PI SETH LAWRENCE SCHOR, ANA MARIA SCHOR
PC C12N15/09, A61K38/00, A61P17/02, C07K14/78, C07K16/18, C12N5/10, PC
C12P21/08,
PC C12Q1/68, G01N33/574, C12N15/00, A61K37/02, C12N5/00 CC
Polypeptides, polynucleotides and uses thereof. FH Key
Location/Qualifiers
FT source 1..2147
/organism="Homo sapiens (human)"
FEATURES
source 1..2147
Location/Qualifiers

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2004, 21:47:11 ; Search time 169.709 Seconds
(without alignments)
3278.783 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQQMVQPSPVAVSQSPG.....SMWDTCTCIGAGRGISCTI 106

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US0940235/runat_03112004_174038_11244/app_query.fasta_1.1045
-DB=N Geneseq_23Sep04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALFT=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @CGN_1_1063 @runat_03112004_174038_11244 -NCFU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_23Sep04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	100.0	777	3 AAA37632	Aaa37632 Human fib
2	600	100.0	1661	3 AAA37637	Aaa37637 Chimeric
3	600	100.0	4860	3 AAA35009	Aaa35009 Human ade
4	600	100.0	6988	9 ACD06169	AcD06169 Human cdn
5	600	100.0	7049	12 ADB95947	Adn95947 Human NOV
6	600	100.0	7679	10 ADB31322	Adb31322 Testoster

RESULT 1

AAA37632	AAA37632 standard; DNA; 777 BP.	
ID	AAA37632 standard; DNA; 777 BP.	
XX		
AC	AAA37632;	
XX		
DT	13-OCT-2000 (first entry)	
XX		
DE	Human fibronectin coding sequence fragment.	
XX		
KW	Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;	
KW	plasminogen; human; fibronectin; thrombolytic therapy;	
KW	cardiovascular disorder; fibronectin; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..777
FT		/*tag= a
FT		/product= "fibronectin"
FT		/partial
XX		
PN	EP1024192-A2.	
XX		
PD	02-AUG-2000.	
XX		
PF	23-DEC-1999;	99EP-00310541.
XX		
PR	24-DEC-1998;	98IN-DE003825.
XX		

ALIGNMENTS

7	600	100.0	7680	2 AAT17551	Aat17551 Human fib
8	600	100.0	7680	3 AAF21131	Aaf21131 Human low
9	600	100.0	7680	5 ABA82689	Aba82689 Fibronect
10	600	100.0	7680	6 ABL67540	AbL67540 Thyroid c
11	600	100.0	7680	6 ABL11082	Abt11082 Human bre
12	600	100.0	7680	8 ACF03878	Acf03878 Human fib
13	600	100.0	7680	8 ABX10391	ABX10391 DNA encod
14	600	100.0	7680	8 ACC46009	Acc46009 Human fib
15	600	100.0	7680	8 ACF12859	Acf12859 Human cer
16	600	100.0	7680	8 ACA64817	Aca64817 Human fib
17	600	100.0	7680	8 ACA64819	Aca64819 Human fib
18	600	100.0	7680	9 ADB70377	Adb70377 Fibronect
19	600	100.0	7680	10 ADB98703	Adb98703 Human fib
20	600	100.0	7680	10 ADD18771	Add18771 Human dis
21	600	100.0	7680	10 AD82499	Ad82499 Human DNA
22	600	100.0	7680	10 ABZ96825	Abz96825 Human nuc
23	600	100.0	7680	11 ADI31963	Adi31963 Human cdn
24	600	100.0	7680	11 ABD20674	Abd20674 Human pul
25	600	100.0	7680	12 ADJ37156	Adj37156 Human mal
26	600	100.0	7680	12 ADP13486	Adp13486 Renal cel
27	600	100.0	7705	1 AAN70596	Aan70596 Sequence
28	600	100.0	7705	2 AAQ15214	Aaq15214 Human fib
29	600	100.0	7705	12 ADG89561	Adg89561 Human fib
30	600	100.0	7803	2 AAQ70009	Aaq70009 Fibrinoge
31	600	100.0	8216	5 AAS86466	Aas86466 DNA encod
32	600	100.0	14740	3 AAA35010	Aaa35010 Human ade
33	600	100.0	14740	3 AAF21132	Aaf21132 Human low
34	600	100.0	14740	10 ABZ96826	Abz96826 Human nuc
35	600	100.0	14741	11 ABD20675	Abd20675 Human pul
36	597	99.5	3540	12 ADP75954	Adp75954 Human min
37	597	99.5	4080	12 ADP75958	Adp75958 Human leu
38	596	99.3	867	3 AAF18086	Aaf18086 Lung canc
39	596	99.3	2127	10 ADD18477	Add18477 Human pro
40	596	99.3	2147	2 AAX81299	Aax81299 Human mig
41	596	99.3	3522	5 AAS86462	Aas86462 DNA encod
42	596	99.3	4295	8 ACC72037	Acc72037 BC00770 g
43	596	99.3	6816	10 ABX74443	Abx74443 Human cdn
44	596	99.3	7361	9 ACD06170	AcD06170 Human cdn
45	596	99.3	7550	8 ACC00412	Acc00412 Human cel

PA (COUL) CSIR COUNCIL SCI IND RES.
XX Sahn G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX WPI; 2000-516032/47.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX domains of human fibronectin.
XX Example 3; Fig 6; 58pp; English.
XX This sequence represents a human fibronectin coding sequence fragment,
XX containing fibrin binding domains. The invention relates to a hybrid
XX plasminogen activator (PA) comprises a polypeptide fusion between
XX streptokinase (SK), which are capable of plasminogen (PG) activation, and
XX fibrin binding regions of human fibronectin, which are from fibrin
XX binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
XX ability to bind with fibrin independently and also characteristically
XX retains a PG activation ability which becomes evident only after a
XX pronounced duration, or lag, after exposure of the PA to a suitable
XX animal or human PG. The hybrid streptokinase-fibrin binding domain
XX polypeptides are useful in thrombolytic therapy for various kinds of
XX cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
XX well as kinetics of plasminogen activation that are distinct from that of
XX natural streptokinase in being characterised by a temporary delay, or lag
XX of several minutes in the natural rate of the catalytic conversion of
XX plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
XX can bind tightly with fibrin in blood clots soon after introduction into
XX the vascular system without significantly activating the circulating
XX blood plasminogen to plasmin, thus aiding in the localisation of the
XX plasminogen activation process to the site of pathological thrombus. This
XX overcomes systemic plasminogen activation encountered during clinical use
XX of streptokinase
XX
SQ Sequence 777 BP; 214 A; 166 C; 230 G; 167 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.31e-60 Length: 777
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x AAA37632 (1-777)
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 1 CAGGCTCAGCAATGGTTTCAGCCCGGTCCTCCCGGTGGCTGTCAGTCAAGCAAGCCCGT 60
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 61 TGTATGACATGGAACACATCATGATATATCAACATGGAGGAGGACCTACCTAGT 120
Qy 41 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 121 AATGTGTTGGTTGTACTTGTATGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 180
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 181 GAAGCTGAAGAGACTTGTCTTTTACCAAGTACCTCGGGAACACATTACCGAGTGGGTGACCT 240
Qy 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 100
Db 241 TATGAGCTCTCTTAAGACTTCCATGATCTGGAGTCTGACTGTGATCGGGGCTGGCGAGG 300
Qy 101 ArgIleSerCysThrIle 106
Db 301 AGAATAAGCTGTACCATC 318

RESULT 2

AAA37637
ID AAA37637 standard; DNA; 1661 BP.
XX
AC AAA37637;
XX
DT 15-SEP-2003 (revised)
DT 13-OCT-2000 (first entry)
XX
DE Chimeric SK-FBD coding sequence.
XX
KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; ss.
XX
OS Streptococcus dysgalactiae subsp. equisimilis.
OS Homo sapiens.
OS Chimeric.
XX
PN EP1024192-A2.
XX
PD 02-AUG-2000.
XX
PF 23-DEC-1999; 99EP-00310541.
XX
PR 24-DEC-1998; 98IN-DE003825.
XX
PA (COUL) CSIR COUNCIL SCI IND RES.
XX
PI Sahn G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
XX WPI; 2000-516032/47.
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX domains of human fibronectin.
XX
PS Example 5; Fig 19b; 58pp; English.
XX
CC This sequence represents a chimeric streptokinase-fibrin binding domain
CC (SK-FBD) protein coding sequence. The invention relates to a hybrid
CC plasminogen activator (PA) comprises a polypeptide fusion between
CC streptokinase (SK), which are capable of plasminogen (PG) activation, and
CC fibrin binding regions of human fibronectin, which are from fibrin
CC binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
CC ability to bind with fibrin independently and also characteristically
CC retains a PG activation ability which becomes evident only after a
CC pronounced duration, or lag, after exposure of the PA to a suitable
CC animal or human PG. The hybrid streptokinase-fibrin binding domain
CC polypeptides are useful in thrombolytic therapy for various kinds of
CC cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
CC well as kinetics of plasminogen activation that are distinct from that of
CC natural streptokinase in being characterised by a temporary delay, or lag
CC of several minutes in the natural rate of the catalytic conversion of
CC plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
CC can bind tightly with fibrin in blood clots soon after introduction into
CC the vascular system without significantly activating the circulating
CC blood plasminogen to plasmin, thus aiding in the localisation of the
CC plasminogen activation process to the site of pathological thrombus. This
CC overcomes systemic plasminogen activation encountered during clinical use
CC of streptokinase
XX
SQ Sequence 1661 BP; 529 A; 370 C; 357 G; 405 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.43e-60 Length: 1661
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x AAA37637 (1-1661)

PF 06-SEP-2002; 2002US-00236392.
 XX 07-SEP-2001; 2001US-0318120P.
 PR 07-SEP-2001; 2001US-0318130P.
 PR 07-SEP-2001; 2001US-0318219P.
 PR 10-SEP-2001; 2001US-0318430P.
 PR 12-SEP-2001; 2001US-0318765P.
 PR 17-SEP-2001; 2001US-0322781P.
 PR 17-SEP-2001; 2001US-0322816P.
 PR 19-SEP-2001; 2001US-0323519P.
 PR 20-SEP-2001; 2001US-0323631P.
 PR 20-SEP-2001; 2001US-0323636P.
 PR 25-SEP-2001; 2001US-0324969P.
 PR 25-SEP-2001; 2001US-0325091P.
 PR 26-SEP-2001; 2001US-0324990P.
 PR 15-FEB-2002; 2002US-0357303P.
 PR 28-FEB-2002; 2002US-0360973P.
 PR 20-MAR-2002; 2002US-0366131P.
 PR 25-MAR-2002; 2002US-0367753P.
 PR 02-APR-2002; 2002US-0369479P.
 PR 10-MAY-2002; 2002US-0379532P.
 PR 17-MAY-2002; 2002US-0381664P.
 PR 17-MAY-2002; 2002US-0381672P.
 PR 28-MAY-2002; 2002US-0383651P.
 PR 29-MAY-2002; 2002US-0384012P.
 PR 19-JUN-2002; 2002US-0390155P.
 XX (ZHON//) ZHONG M.
 PA (LILL//) LI L.
 PA (GORM//) GORMAN L.
 PA (SPYT//) SPYTEK K A.
 PA (KEKU//) KEKUDA R.
 PA (TAUP//) TAUPIER R J.
 PA (ANDE//) ANDERSON D W.
 PA (VERN//) VERNET C A M.
 PA (CATT//) CATTERTON E.
 PA (MILL//) MILLER C E.
 PA (SHEN//) SHENOV S G.
 PA (PATT//) PATTURAJAN M.
 PA (PENA//) PENA C E A.
 PA (TCHER//) TCHERNEV V T.
 PA (PADI//) PADIGARU M.
 PA (GUSE//) GUSEV V Y.
 PA (MALY//) MALYANKAR U M.
 PA (BURG//) BURGESS C E.
 PA (GERL//) GERLACH V.
 PA (CASM//) CASMAN S J.
 PA (RIEG//) RIEGER D K.
 PA (GROS//) GROSSE W M.
 PA (SMIT//) SMITHSON G.
 PA (PEYM//) PEYMAN J A.
 PA (STAR//) STARLING G.
 PA (ROTH//) ROTHENBERG M E.
 PA (LARO//) LAROCHELLE W J.
 PA (SHIM//) SHIMKETS R A.
 PA (CRAB//) CRABTREE J.
 PA (RAST//) RASTELLI L.
 PA (VOSS//) VOSS E Z.
 PA (BOLD//) BOLDOG F L.
 PA (EDIN//) EDINGER S R.
 PA (MILL//) MILLET I.
 PA (MACD//) MACDOUGALL J R.
 PA (ELLE//) ELLERMAN K.
 PA (CHAP//) CHAPOVAL A.
 XX Zhong M, Li L, Gorman L, Spytek KA, Kekuda R, Taupier RJ;
 PI Anderson DW, Vernet CAM, Catterton E, Miller CE, Shenov SG;
 PI Patturajan M, Pena CE, Tchernev VT, Padigar M, Gusev VY;
 PI Malyankar UM, Burgess CE, Gerlach V, Casman SJ, Rieger DK;
 PI Larochelle WJ, Smithson G, Peyman JA, Starling G, Rothenberg ME;
 PI Boldog FL, Edinger SR, Millet I, Macdougall JR, Ellerman K;
 PI Chapoval A;

XX WPI; 2004-355290/33.
 DR P-PSDB; ADN95948.
 XX New isolated polypeptide, useful for treating or preventing a pathology associated with the polypeptide, e.g. diabetes, infectious disease, cancer, neurodegenerative disorders or Alzheimer's disease.
 PT
 PT
 XX
 PS Claim 20; SEQ ID NO 1; 552pp; English.
 XX
 CC The invention relates to human NOVX polypeptides and polynucleotides. The isolated nucleic acids can be used to express the novel proteins, to detect novel mRNA or a genetic lesion in a novel gene and to modulate its activity. It can also be used in gene therapy for treating or preventing a pathology associated with the protein or nucleic acid. The disorders include metabolic disorders, diabetes, obesity, infectious diseases, anorexia, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders and haematopoietic disorders. This sequence represents a human NOVX polynucleotide of the invention.
 CC
 XX
 SQ Sequence 7049 BP; 1888 A; 1972 C; 1651 G; 1535 T; 0 U; 3 Other;

Alignment Scores:
 Pred. No.: 2.13e-59 Length: 7049
 Score: 600.00 Matches: 106
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 12 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x ADN95947 (1-7049)
 Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
 Db 14 CAGGCTCAGCAAAATGGTTTCAGCCCGAGTCCCGTGGCTGTCTAGTCAAGCAAGCCCGGT 73
 Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
 Db 74 TGTATTGACAATCGAAACACATATCAGATAAATCAACAGTGGGAGCGACCTACCTAGGT 133
 Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
 Db 134 AATGCTGTGGTTTGTACTTCTTATGGAGGAAGCGAGGTTTAACTCGAAAGTAAACCT 193
 Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
 Db 194 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 253
 Qy 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 100
 Db 254 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTCTGCATCGGGGCTGGCGAGGG 313
 Qy 101 ArgIleSerCysThrIle 106
 Db 314 AGAATAAGCTGTACCATC 331
 RESULT 6
 ADB31322
 ID ADB31322 standard; cDNA; 7679 BP.
 XX
 AC ADB31322;
 XX
 XX 04-DEC-2003 (first entry)
 DT
 XX
 DE Testosterone regulated prostate cancer gene #38.
 XX Human; sg; prostate cancer; cancer; androgen; castration;
 KW anti-androgenic drug; bicalutamide; casodex; testosterone.
 XX Homo sapiens.
 OS
 XX US6506607-B1.
 PN
 XX

PD 14-JAN-2003.
XX
XX 23-DEC-1998; 98US-00220132.
XX
XX 24-DEC-1997; 97US-0068821P.
PR 25-MAR-1998; 98US-0079303P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Shyjan AW;
XX
XX WPI; 2003-595981/56.
XX
XX Determination of whether compound is useful for prostate cancer treatment
PT comprises measuring expression level of specific nucleic acid sequence in
PT prostate cancer cell sample in the presence and absence of compound.
XX
XX Example 1; Col 139-146; 194pp; English.
XX
XX The invention discloses a method for determining whether a compound can
CC be used to treat prostate cancer and comprises measuring the expression
CC level of a nucleic acid in prostate cancer cell sample in the presence
CC and absence of the compound. Determining whether a compound can be used
CC to treat prostate cancer comprises identifying the compound as useful for
CC prostate cancer treatment when the expression level of the nucleic acid
CC in the presence of the compound is less than the expression level of the
CC nucleic acid in the absence of the compound. Prostate cancer is usually
CC treated by androgen withdrawal, by castration or through the use of an
CC anti-androgenic drug. Bicalutamide (casodex) is one such anti-androgenic
CC compound. The determination whether a selected compound, e.g. an anti-
CC androgenic compound or testosterone, can be used to treat the prostate
CC cancer can be made on a patient by patient basis. The sequence presented
CC is a gene which is more highly expressed in testosterone treated prostate
CC cancer cells than the untreated cells.
XX
SQ Sequence 7679 BP; 2081 A; 1981 C; 1862 G; 1755 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,378-59 Length: 7679
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x ADB31322 (1-7679)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAATGGTTTCAGCCCAAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATTGACAAATGGAAACACTATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
DB 134 AATGTGTTGGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGCAAGTAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGAACAAGTACACTTGGGAACACTTACCAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCCCTAAAGACTCCATGATCTCGGACTGTACCTGCACTCGGGCTGGGCGAGGG 313
QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331

RESULT 7
AAT17551

ID AAT17551 standard; cDNA; 7680 BP.
XX
XX AAT17551;
XX
DT 21-JUN-1996 (first entry)
XX
XX Human fibronectin cDNA.
XX
XX Fibrin-binding peptide; fibronectin; fibrinolysis; fibrinogenesis;
KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis; therapy;
KW ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT 14..6988
FT /*tag= a
XX
XX WO9604304-A1.
XX
XX 15-FEB-1996.
XX
XX 01-AUG-1995; 95WO-US009819.
XX
XX 01-AUG-1994; 94US-00283857.
XX
XX (UUNY) UNIV NEW YORK STATE.
XX
XX Gold LI, Rostagno AA;
PI
XX WPI; 1996-129333/13.
DR
XX P-PSDB; AAR92778.
XX
XX New fibrin-binding peptide molecules - used for the diagnosis and
PT treatment of conditions associated with fibrin deposition, e.g. thrombi.
XX
XX Disclosure; Page 112-115; 146pp; English.
XX
XX A cDNA sequence (AAT17551) codes for human fibronectin (AAR92778). It is
CC used for the prodn. of fibrin-binding modules of fibronectin useful in
CC the diagnosis and treatment of conditions associated with fibrin
CC deposition, e.g. for disruption of blood clots
XX
XX Sequence 7680 BP; 2083 A; 1979 C; 1862 G; 1756 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,378-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x AAT17551 (1-7680)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAATGGTTTCAGCCCAAGTCCCGGTGGCTGTCAAGTCAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATTGACAAATGGAAACACTATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
DB 134 AATGTGTTGGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGCAAGTAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGAACAAGTACACTTGGGAACACTTACCAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100

Db 254 TATGAGCGTCTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG 313
Qy 101 ArgileSerCysThrIle 106
|||||
Db 314 AGAATAAGCTGTACCATC 331
RESULT 8
AAF21131
ID AAF21131 standard; DNA; 7680 BP.
XX
XX AAF21131;
XX
XX 14-MAR-2001 (first entry)
XX
XX Human low adenosine antisense oligonucleotide related sequence #2698.
XX
XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200062736-A2.
XX
XX 26-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US008020.
XX
XX 06-APR-1999; 99US-0127958P.
XX
XX (UYEC-) UNIV EAST CAROLINA.
XX
XX (NYCE/) NYCE J W.
XX
XX Nyce JW;
XX
XX WPI; 2000-679539/66.
XX
XX Low adenosine (A) content antisense oligonucleotides which do not trigger
PT adenosine receptors during metabolism, useful e.g. for treating cancers
PT and respiratory obstructions.
XX
XX Disclosure; Page 1026-1028; 1592pp; English.
XX
XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or
CC surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome

CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention
XX
SQ Sequence 7680 BP; 2081 A; 1981 C; 1962 G; 1756 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,37e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x AAF21131 (1-7680)
Qy 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGGTTTCAGCCCAAGTCCCGGTGGCTGTCTCAGTCAAGCAAGCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATTGACAATGGAAAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTGTGGTTGGTTGTACTTGTATGGAGGAAGCGGAGTTTAACTCGGAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAGAGACTTGTCTTTGACAAGTACACTGGGAAACACTTACCAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG 313
Qy 101 ArgileSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331
RESULT 9
ABA82689
ID ABA82689 standard; DNA; 7680 BP.
XX
XX ABA82689;
XX
XX 25-JAN-2002 (first entry)
XX
XX Fibronectin gene SEQ ID NO:75.
XX
XX Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
KW sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
KW antisense therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
KW osteomalacia; fibrous dysplasia; ds.
XX
XX Homo sapiens.
XX
XX WO200177327-A1.
XX
XX 18-OCT-2001.
XX
XX 21-JUN-2000; 2000WO-US016951.
XX
XX 05-APR-2000; 2000US-00543771.
XX
XX 05-APR-2000; 2000US-00544398.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Carulli JP, Little RD, Recker RR, Johnson ML;
XX

DR WPI; 2001-657171/75.
XX New high bone mass (HBM) and Zmax1 genes and proteins useful for
PT modulating bone mass for the treatment of e.g. osteoporosis.
XX
PS Claim 79; Page 376-378; 443pp; English.
XX
CC The present invention describes the human Zmax1 gene and the high bone
CC mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
CC genes have osteopathic activities. The genes can be used in gene therapy,
CC antineuse therapy and in the production of vaccines. They can be used in
CC the diagnosis and treatment of bone disorders including osteoporosis,
CC Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
CC ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2, 37e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x ABA82689 (1-7680)
Oy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerIysProGly 20
Db 14 CAGGCTCAGCAATGTTTCAGCCCGTCCCGTGTGTCAGTCAAGCAAGCCCGGT 73
Oy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTyrGluArgThrTyrIleuGly 40
Db 74 TGTATGACATATGGAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
Oy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerIysPro 60
Db 134 AATGTTGGTTTGTACTTGTATGGAGGAGCGGAGTTTAACTGCGAAGTAACT 193
Oy 61 GluAlaGluGluThrCysPheAspIysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTTCACAAAGTACACTGGGAAACACTTACCGAGTGGTGACACT 253
Oy 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGCTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGCTGGCGAGGG 313
Oy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAGCTGTACCATC 331

RESULT 10
ABL67540
ID ABL67540 standard; DNA; 7680 BP.
XX
AC ABL67540;
XX
DT 15-MAY-2002 (first entry)
XX
DE Thyroid cancer related gene sequence SEQ ID NO:5877.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytosstatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
FN WO200194629-A2.
XX
PD 13-DEC-2001.
XX

PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 22-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 5877; 4pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S); where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous

CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
XX tumour
SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,37e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x ABL67540 (1-7680)
QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAATGTTTGTGAGCCCGCCAGTCCCGGTGGCTGTCTAGTCAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
DB 134 AATGTGTGTGGTTTGTACTTGTATGGAGGAGCGGAGGTTTAACTCGCAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCAGTGGGTGACCT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGCGTGGGCTGGGCGAGGG 313
QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331
RESULT 11
ABT11082
ID ABT11082 standard; cDNA; 7680 BP.
AC
XX
AC ABT11082;
XX
XX
DT 04-DEC-2002 (first entry)
XX
XX Human breast cancer associated coding sequence SEQ ID NO: 1216.
DE
XX Human; breast specific gene; breast cancer; differential expression;
KW cyrostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200259271-A2.
XX
XX
PD 01-AUG-2002.
XX
XX 25-JAN-2002; 2002WO-US002176.
XX
XX 25-JAN-2001; 2001US-0263757P.
PR 25-APR-2001; 2001US-0286090P.
PR 23-MAY-2001; 2001US-0292517P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Orr MS, Nation M, Diggins JC, Zeng W;
XX
XX WPI; 2002-674803/72.
DR
XX Diagnosing breast cancer in a patient comprises detecting the level of
PT gene expression in cell or tissue samples, where a differential gene
PT expression is indicative of breast cancer.

XX Claim 1; SEQ ID NO 1216; 260pp + Sequence Listing; English.
XX
CC The present invention relates to methods of diagnosing breast cancer in a
CC patient, which comprise detecting the level of expression in a tissue
CC sample of two or more genes selected from those shown in ABT09867-
CC ABT1112, where a differential expression of the genes indicates breast
CC cancer. The methods are useful in diagnosing, treating, detecting the
CC progression, and in monitoring treatment of breast cancer in patients.
CC The methods are also useful as a screening tool for agents that modulate
CC the onset or progression of breast cancer. The breast cancer genes may be
CC used as diagnostic markers for the prediction or identification of the
CC malignant state of breast tissue, for confirming the type and progression
CC of cancer, and for drug screening and assays. The present sequence is a
CC coding sequence of the invention. Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub.published_pct_sequences
XX
SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,37e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x ABT11082 (1-7680)
QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAATGTTTGTGAGCCCGCCAGTCCCGGTGGCTGTCTAGTCAAGCAAGCCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATTGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
DB 134 AATGTGTGTGGTTTGTACTTGTATGGAGGAGCGGAGGTTTAACTCGCAAGTAAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTGGGAACACTTACCAGTGGGTGACCT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGCGTGGGCTGGGCGAGGG 313
QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331
RESULT 12
ACF03878
ID ACF03878 standard; cDNA; 7680 BP.
XX
XX ACF03878;
XX
DT 18-SEP-2003 (first entry)
XX
XX Human fibronectin encoding cDNA SEQ ID NO:9.
DE
XX Human; integrin alpha4beta1; fibronectin; integrin alpha 4; VCAM;
KW integrin beta 1; vascular cell adhesion molecule; angiogenesis;
KW integrin alpha4beta1 binding inhibitor; angiogenesis inhibitor;
KW cyrostatic; ophthalmological; antirheumatic; antiarthritic; cancer;
KW antiinflammatory; osteopathic; dermatological; gene therapy; psoriasis;
KW pathological condition; diabetic retinopathy; macular degeneration;
KW neovascularisation; rheumatoid arthritis; osteoarthritis; skin cancer;
KW endothelial progenitor cell; wound healing; gene; ss.

XX OS Homo sapiens.
XX WO2003019136-A2.
XX PD 06-MAR-2003.
XX PF 01-AUG-2002; 2002WO-US024573.
XX PR 06-AUG-2001; 2001US-0310645P.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Varner JA;
XX DR WPI; 2003-278690/27.
XX DR P-PSDB; ABR81866.
XX
XX Inhibiting angiogenesis in a tissue, useful for treating cancer,
PT arthritis, retinopathy, psoriasis, by providing a tissue and an agent
PT that inhibits specific binding of integrin alpha-4beta-1 to an integrin
PT alpha-4beta-1 ligand.
XX
XX Disclosure; Fig 17; 177pp; English.
XX
XX The present invention describes a method for inhibiting angiogenesis in a
CC tissue comprising providing a tissue and an agent that inhibits specific
CC binding of integrin alpha4beta1 to an integrin alpha4beta1 ligand, and
CC treating the tissue with the agent under conditions where the specific
CC binding of the alpha4beta1 to the ligand is inhibited and a treated
CC tissue is produced. Angiogenesis in the treated tissue is inhibited. Also
CC described: (1) inhibiting endothelial cell adhesion or migration; (2)
CC detecting angiogenesis in a tissue; (3) screening a test compound; (4)
CC isolating endothelial progenitor cells from a tissue; and (5) reducing
CC symptoms associated with cancer in a subject, or a pathological condition
CC is an ocular or skin tissue. An integrin alpha4beta1 binding inhibitor
CC has cytostatic, ophthalmological, antirheumatic, antiarthritic,
CC antiinflammatory, osteopathic and dermatological activities, and can be
CC used in gene therapy. The methods are useful for treating cancer, and
CC other pathological condition, such as diabetic retinopathy, macular
CC degeneration by neovascularisation, rheumatoid arthritis, osteoarthritis,
CC psoriasis or skin cancer. The methods are also useful in isolating
CC endothelial progenitor cells, and in determining the mechanisms that
CC underlie angiogenesis, development, wound healing and the function of the
CC female reproductive system. The present sequence is used in the
CC exemplification of the present invention
XX
XX Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.37e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x ACF03878 (1-7680)

Qy 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGTTTGTGAGCCCGCCAGTCCCGGGCTGTGAGTCAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACATGGAACACATATCAGATAAATCAACAGCTGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPhaAsnCysGluSerLysPro 60
Db 134 AATGTGTTGGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGCAAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80

Db 194 GAAGCTGAAGAGACTTGTCTTTGACAAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTtpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACTGCAATCGGGCTGGCGGAGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331
RESULT 13
ID ABX10391
ID ABX10391 standard; DNA; 7680 BP.
XX AC ABX10391;
XX DT 28-JAN-2003 (first entry)
XX DE DNA encoding protein differentially regulated in prostate cancer #60.
XX KW Prostate cancer; gene expression; differential regulation;
KW molecular marker; drug target; cancer detection; cancer diagnosis;
KW cancer staging; cancer grading; cancer assessing; cancer monitoring;
KW gene; ds.
XX OS Homo sapiens.
XX PN WO200281638-A2.
XX PD 17-OCT-2002.
XX PF 08-APR-2002; 2002WO-US010824.
XX PR 06-APR-2001; 2001US-0281731P.
XX PR 06-APR-2001; 2001US-0281732P.
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX PI Sun Z, Jay G;
XX WPI; 2003-058520/05.
XX P-PSDB; ABU07486.
XX Novel genes which are differentially regulated in prostate cancer, useful
PT for diagnosing prostate cancer in prostate tissue sample and assessing
PT therapeutic or preventive intervention in prostate cancer patients.
XX
XX Claim 1; Page 136-139; 416pp; English.
XX
XX The invention describes genes (I) which are differentially regulated in
CC prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC sample comprising prostate tissue, which involves determining the number
CC of target genes which are differentially-regulated in the sample, where
CC the number is indicative of the probability that the sample comprises
CC prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC intervention in a subject having a prostate cancer, which involves
CC determining the expression levels in a sample comprising prostate tissue
CC of target genes which are differentially-regulated in prostate cancer.
CC Preferably, the expression levels of at least 10 genes are determined.
CC (I) is also useful for identifying agents that modulate a biological
CC activity of a polypeptide differentially-regulated in prostate cancer
CC cells, which involves contacting a polypeptide differentially-regulated
CC in prostate cancer cells with a test agent under conditions effective for
CC the test agent to modulate a biological activity of the polypeptide, and
CC determining whether the test agent modulates the biological activity. (I)
CC is useful as molecular markers, as drug targets, and for detecting,
CC diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC preventing or treating, determining predisposition to diseases and
CC conditions especially relating to prostate cancer. (I) and its expression
CC products are used in the diagnostic test to assay for presence of cancer
CC e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC of cancer, its stage of development, the nature of genetic defect, etc.

RESULT 15
ACF12859
ID ACF12859 standard; cDNA; 7680 BP.
XX
AC ACF12859;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human cervical cancer cell marker encoding cDNA SEQ ID NO:63.
XX
KW Human; cervical cancer; cervical cancer marker; cancer therapy;
XX detection; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2002101075-A2.
XX
PD 19-DEC-2002.
XX
PF 12-JUN-2002; 2002WO-US018638.
XX
PR 13-JUN-2001; 2001US-0298155P.
XX
PR 13-JUN-2001; 2001US-0298159P.
XX
PR 14-NOV-2001; 2001US-0335936P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;
XX
DR P-PSDB; ABR92078.
XX
WPI: 2003-156967/15.
XX
P-PSDB; ABR92078.
XX
New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
Claim 4; Page 209-212; 386pp; English.
XX
ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (MI) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials
XX
SQ Sequence 7680 BP; 2081 A; 1981 C; 1862 G; 1756 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.37e-59 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x ACF12859 (1-7680)
Qy 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGTTTTCAGCCCGGTCAGTCCAGTCAAAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40

Db 74 TGTATTGACAAATGGGAAAAACACTATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTTGTTGTTTGTACTTGTATGAGGAGGAGGTTTAACTGCGAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAAGAGACTTGTCTTTGACAAGTACACTGGGGAACACTTACCGAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTtpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

Search completed: November 6, 2004, 00:37:53
Job time : 180.709 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2004, 00:01:06 ; Search time 1295.05 Seconds
(without alignments)
2982.593 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQMVPQSPVAVSQSPKPG.....SMWDCITGAGRGRISCTI 106

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_03112004_174039_11264/app_query.fasta_1.1045
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 -CGN_1_1_8076/runat_03112004_174039_11264 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	596	99.3	480	4 BF993488	BF993488 QVO-GN021
2	596	99.3	487	4 BF993485	BF993485 QVO-GN021
3	596	99.3	597	7 CN332340	CN332340 170005325
4	596	99.3	657	7 CN332343	CN332343 170005314
5	596	99.3	660	7 CN332334	CN332334 170004554
6	596	99.3	666	7 CN332338	CN332338 170006001
7	596	99.3	673	7 CN419479	CN419479 170004706
8	596	99.3	676	1 AL603362	AL603362 DKFZp686C
9	596	99.3	684	7 CN332332	CN332332 170005326

10	596	99.3	706	7	CN332335	CN332335 170005322
11	596	99.3	716	7	CN332342	CN332342 170005339
12	596	99.3	751	1	AL706288	AL706288 DKFZp686P
13	596	99.3	756	5	EX473407	EX473407 DKFZp686G
14	596	99.3	765	5	EX473413	EX473413 DKFZp686I
15	596	99.3	911	5	EX398838	EX398838 BX398838
16	592	98.7	676	5	EX510005	EX510005 DKFZp686C
17	592	98.7	748	5	EX473391	EX473391 DKFZp686C
18	592	98.7	913	5	EX380583	EX380583 BX380583
19	591	98.5	693	1	AL706197	AL706197 DKFZp686I
20	591	98.5	699	1	AL706221	AL706221 DKFZp686J
21	589	98.2	427	4	BF988908	BF988908 IL5-GN017
22	589	98.2	689	5	EX473425	EX473425 DKFZp686L
23	589	98.2	719	5	EX473443	EX473443 DKFZp686P
24	589	98.2	749	5	EX473395	EX473395 DKFZp686D
25	589	98.2	759	5	EX473392	EX473392 DKFZp686C
26	588	98.0	434	5	BQ366579	BQ366579 QVO-GN021
27	588	98.0	683	5	EX501294	EX501294 DKFZp779A
28	587	97.8	1026	5	EX386270	EX386270 BX386270
29	587	97.8	1038	5	EX439175	EX439175 BX439175
30	587	97.5	1050	5	EX417945	EX417945 BX417945
31	585	97.5	378	2	BF751799	BF751799 RC3-BN042
32	583	97.2	321	2	AW068245	AW068245 cn23907.Y
33	583	97.2	684	7	CN332337	CN332337 170005326
34	583	97.2	712	1	AL706173	AL706173 DKFZp686G
35	580	96.7	581	6	CD544727	CD544727 B0260C05-
36	580	96.7	594	7	CF909456	CF909456 A0536G09-
37	580	96.7	605	6	BY742468	BY742468 BY742468
38	580	96.7	611	7	CF907410	CF907410 A0510811-
39	580	96.7	622	2	BB866384	BB866384 BB866384
40	580	96.7	623	2	BB866390	BB866390 BB866390
41	580	96.7	628	2	BB865942	BB865942 BB865942
42	580	96.7	635	2	BB866114	BB866114 BB866114
43	580	96.7	637	2	BB866721	BB866721 BB866721
44	580	96.7	637	2	BB865910	BB865910 BB865910
45	580	96.7	637	2	BB866224	BB866224 BB866224

ALIGNMENTS

RESULT 1

BF993488

LOCUS

QVO-GN0216-011100-466-g03

GN0216 Homo sapiens

linear

EST

BF993488.1 GI:12399811

EST.

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 480)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., and

O'Hare,W.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

DEFINITION 17000532545117 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN3323340
 VERSION CN332340.1 GI:47332274
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 597)
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J and Stanton, L.W.
 TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
 JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
 COMMENT Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 597 Std Error: 0.00.
 FEATURES
 source
 1..597
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
 /clone_lib="GRN_EB"
 note="oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from HES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 89e-61 Length: 597
 Score: 596.00 Matches: 105
 Percent Similarity: 99.06% Conservative: 0
 Best Local Similarity: 99.06% Mismatches: 1
 Query Match: 99.33% Indels: 0
 DB: 7 Gaps: 0
 US-09-940-235-4_COPY_1_106 (1-106) x CN332340 (1-597)
 QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
 DB 257 CAGGCTCAGCAATGGTTGAGCCCAAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCGGT 316
 QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
 DB 317 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGACCTACCTAGGC 376
 QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
 DB 377 AATGCGTGGTTGTACTTGTATGGAGAACCGAGGTTTAACTCGCAGAGTAACCT 436
 QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
 DB 437 GAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 496
 QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
 DB 497 TATGAGCGTCTTAAGACTCCATGATCTGGACTGTACCTGCTACCTCGGGCTGGCGGAGG 556
 QY 101 ArgIleSerCysThrIle 106
 DB 557 AGAATAAGCTGTACCATC 574
 RESULT 4
 CN332343

LOCUS CN332343
 DEFINITION 17000531483454 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN332343
 VERSION CN332343.1 GI:47332277
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 657)
 AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J and Stanton, L.W.
 TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
 JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
 COMMENT Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 657 Std Error: 0.00.
 FEATURES
 source
 1..657
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, embryoid bodies derived from H1, H7 and H9 cells"
 /clone_lib="GRN_EB"
 note="oligo dt primed, full-length enriched cDNA library from embryoid body outgrowths derived from HES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."
 ORIGIN
 Alignment Scores:
 Pred. No.: 4 39e-61 Length: 657
 Score: 596.00 Matches: 105
 Percent Similarity: 99.06% Conservative: 0
 Best Local Similarity: 99.06% Mismatches: 1
 Query Match: 99.33% Indels: 0
 DB: 7 Gaps: 0
 US-09-940-235-4_COPY_1_106 (1-106) x CN332343 (1-657)
 QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
 DB 289 CAGGCTCAGCAATGGTTGAGCCCAAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCGGT 348
 QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
 DB 349 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGACCTACCTAGGC 408
 QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
 DB 409 AATGCGTGGTTGTACTTGTATGGAGAACCGAGGTTTAACTCGCAGAGTAACCT 468
 QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
 DB 469 GAAGCTGAAGAGACTTGTCTTGAACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 528
 QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
 DB 529 TATGAGCGTCTTAAGACTCCATGATCTGGACTGTACCTGCTACCTCGGGCTGGCGGAGG 588
 QY 101 ArgIleSerCysThrIle 106
 DB 589 AGAATAAGCTGTACCATC 606
 RESULT 5

```
CN332334
LOCUS       CN332334               660 bp mRNA linear EST 16-MAY-2004
DEFINITION  17000455431365 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN332334
VERSION     CN332334.1 GI:47332268
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 660)
AUTHORS     Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE       Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL     Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 660 Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1..660
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cells, embryoid bodies
                     derived from H1, H7 and H9 cells"
                     /clone_lib="GRN_EB"
                     /note="oligo dt primed, full-length enriched cDNA library
                     from embryoid body outgrowths derived from hES cell lines
                     H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
                     conditions."
ORIGIN
Alignment Scores:
Pred. No.:      4,42e-61      Length:      660
Score:          596.00      Matches:    105
Percent Similarity: 99.06%      Conservative: 0
Best Local Similarity: 99.06%      Mismatches: 1
Query Match:    99.33%      Indels:    0
DB:              7      Gaps:      0
US-09-940-235-4_COPY_1_106 (1-106) x CN332334 (1-660)
```

```
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 327 CAGGCTCAGCAAAATGGTTTCAGCCCAAGTCCCGGTGGCTGTTCAGTCAAGCAAGCCCGT 386
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 387 TGTATGACATGGAACACACTATCGATATCAACACTGGAGCGGACCTACCTAGGC 446
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 447 AATGCGTTGGTTTGTACTTGTATGAGGAGCCGAGGTTTAACTCGAGAGTAAACCT 506
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 507 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAACACTTACCAGTGGGTGACACT 566
Qy 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysTleGlyAlaGlyArgGly 100
Db 567 TATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTGCTGATCGATCGGGCTGGCGAGG 626
Qy 101 ArgIleSerCysThrIle 106
Db 627 AGAATAAGCTGTACCATC 644
```

```
RESULT 6
LOCUS       CN332338               666 bp mRNA linear EST 16-MAY-2004
DEFINITION  17000600181351 GRN_PRENEU Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   CN332338
VERSION     CN332338.1 GI:47332272
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 666)
AUTHORS     Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.
TITLE       Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL     Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT     Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 666 Std Error: 0.00.
FEATURES             Location/Qualifiers
     source          1..666
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /tissue_type="embryonic stem cell, retinoic acid and
                     mitogen-treated hES cell line H7"
                     /clone_lib="GRN_PRENEU"
                     /note="oligo dt primed, full-length enriched cDNA library
                     from hES cell line H7 (p29) maintained in feeder-free
                     conditions. Embryoid bodies were generated in the presence
                     of all-trans retinoic and mitogens."
ORIGIN
Alignment Scores:
Pred. No.:      4,47e-61      Length:      666
Score:          596.00      Matches:    105
Percent Similarity: 99.06%      Conservative: 0
Best Local Similarity: 99.06%      Mismatches: 1
Query Match:    99.33%      Indels:    0
DB:              7      Gaps:      0
US-09-940-235-4_COPY_1_106 (1-106) x CN332338 (1-666)
```

```
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 299 CAGGCTCAGCAAAATGGTTTCAGCCCAAGTCCCGGTGGCTGTTCAGTCAAGCAAGCCCGT 358
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 359 TGTATGACATGGAACACACTATCGATATCAACACTGGAGCGGACCTACCTAGGC 418
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 419 AATGCGTTGGTTTGTACTTGTATGAGGAGCCGAGGTTTAACTCGAGAGTAAACCT 478
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 479 GAAGCTGAAGAGACTTGTCTTTGACAAGTACACTTGGGAACACTTACCAGTGGGTGACACT 538
Qy 81 TyrGluArgProLysAspSerMetIleTyrAspCysThrCysTleGlyAlaGlyArgGly 100
Db 539 TATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTGCTGATCGGGCTGGCGAGG 598
Qy 101 ArgIleSerCysThrIle 106
Db 599 AGAATAAGCTGTACCATC 616
```

```

RESULT 7
CN419479
LOCUS      CN419479      673 bp      mRNA      linear      EST 16-MAY-2004
DEFINITION 17000470666007 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION  CN419479
VERSION     CN419479.1 GI:47407073
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 673)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 673 Std Error: 0.00.
Location/Qualifiers
1..673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/note="Oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from HES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN
Alignment Scores:
Pred. No.: 4,53e-61 Length: 673
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
Gaps: 0
DB: 7

US-09-940-235-4_COPY_1_106 (1-106) x CN419479 (1-673)

QY 1 GluAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 195 CAGGCTCAGCAATGTTTCAGCCCACTCCCGTGGCTGCTCAGTCAAGCAAGCCCGT 254
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 255 TGTATTGACAATGGAAAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGC 314
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 315 ANTGGCTTGGTTGTTACTTGTATGGAGGAGCCGAGGTTTAACTCCGAGAGTAACT 374
QY 61 GluAlaGluGlnThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 375 GAAGCTGAAGAGACTTGTCTTGGACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 434
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 435 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGG 494
QY 101 ArgIleSerCysThrIle 106
|||||

```

```

Db 495 AGAATAAGCTGTACCATC 512

RESULT 8
AL603362      676 bp      mRNA      linear      EST 04-SEP-2003
LOCUS      AL603362
DEFINITION DKFZp686C197_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION  AL603362
VERSION     AL603362
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 676)
Ansorge, W., Wirkner, U., Mewes, W., Weil, B. and Wiemann, S.
EST (Ansorge, W., Wirkner, U., Mewes, H.W., Weil, B. and Wiemann, S.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp686C197) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..676
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686C197"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4,56e-61 Length: 676
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
Gaps: 0
DB: 1

US-09-940-235-4_COPY_1_106 (1-106) x AL603362 (1-676)

QY 1 GluAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 20 CAGGCTCAGCAATGTTTCAGCCCACTCCCGTGGCTGCTCAGTCAAGCAAGCCCGT 79
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 80 TGTATTGACAATGGAAAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGC 139
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 140 AATGCGTGTGGTTGTACTTGTATGGAGGAGCCGAGGTTTAACTCCGAGAGTAACT 199
QY 61 GluAlaGluGlnThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 200 GAAGCTGAAGAGACTTGTCTTGGACAAGTACACTGGGAACACTTACCAGTGGGTGACACT 259
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 260 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGG 319

```

```

QY 101 ArgileSerCysThrIle 106
Db 320 AGAATAAGCTGTACCATC 337

RESULT 9
LOCUS CN332332 684 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532625776 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN332332
VERSION CN332332.1 GI:47332266
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 684)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 684 Std Error: 0.00.
FEATURES
source
1..684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."
ORIGIN
Alignment Scores:
Pred. No.: 4,63e-61 Length: 684
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 7 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x CN332332 (1-684)
QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 314 CAGGCTCAGCAATGGTTCAGCCCGGCTGCTCAGTCAAGCAAGCCCGGT 373
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 374 TGTATGACATGGAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGC 433
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db 434 AATGGTTGGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGAGAGTAAACCT 493
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 494 GAAGCTGAAGAGACTTGGCTTTGACAAGTAGTACCTGGGAACACTTACCAGTGGGTGACACT 553
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100

```

```

Db 554 TATGAGCGTCTAAAGACTCCATGATCTGGGACTGTACCTGCTGCGGGCTGGCGAGGG 613
QY 101 ArgileSerCysThrIle 106
Db 614 AGAATAAGCTGTACCATC 631

RESULT 10
LOCUS CN332335 706 bp mRNA linear EST 16-MAY-2004
DEFINITION 17000532295426 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN332335
VERSION CN332335.1 GI:47332269
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W
TITLE Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 706 Std Error: 0.00.
FEATURES
source
1..706
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."
ORIGIN
Alignment Scores:
Pred. No.: 4,82e-61 Length: 706
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 7 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x CN332335 (1-706)
QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 227 CAGGCTCAGCAATGGTTCAGCCCGGCTGCTCAGTCAAGCAAGCCCGGT 286
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 287 TGTATGACATGGAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGC 346
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db 347 AATGGCTTGGTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGAGAGTAAACCT 406
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 407 GAAGCTGAAGAGACTTGGCTTTGACAAGTAGTACCTGGGAACACTTACCAGTGGGTGACACT 466
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100

```

```

Db      467 TATGAGCGTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGCGTGGCGAGGG 526
Qy      101 ArgileSerCysThrIle 106
Db      527 AGAATAAGCTGTACCATC 544

RESULT 11
LOCUS   CN332342
DEFINITION 716 bp mRNA linear EST 16-MAY-2004
ACCESSION 17000533983768 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
VERSION  CN332342.1 GI:47332276
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 716)
AUTHORS  Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W
TITLE  Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
JOURNAL  Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT  Contact: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 716 Std Error: 0.00.

FEATURES             source
    source
    1..716
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /tissue_type="embryonic stem cells, embryoid bodies
        derived from H1, H7 and H9 cells"
        /clone_lib="GRN EB"
        /note="oligo dt primed, full-length enriched cDNA library
        from embryoid body outgrowths derived from hES cell lines
        H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
        conditions."

ORIGIN
Alignment Scores:
Pred. No.:      4 91e-61      Length:      716
Score:          596.00      Matches:      105
Percent Similarity: 99.06%      Conservative: 0
Best Local Similarity: 99.06%      Mismatches: 1
Query Match:     99.33%      Indels:      0
DB:              7          Gaps:      0

US-09-940-235-4_COPY_1_106 (1-106) x CN332342 (1-716)
Qy      1  GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db      237 CAGGCTCAGCAAAATGGTTGAGCCCGGAGTCCCGGCTGCTGCTAGTCAAAAGCAAGCCCGT 296
Qy      21  CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db      297 TGTATGACATATGGAACACATATATCAGATAAAATCAACAGTGGGAGCGGACCTACCTAGGC 356
Qy      41  AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db      357 AATGCGTTGGTTGTACTGTTATGGAGAACCCGAGGTTTAACTGCGAGAGTAAACCT 416
Qy      61  GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db      417 GAAGCTGAAGAGACTTGTCTTGACAAAGTACATCGGGAACACTTACCGAGTGGGTGACACT 476

```

```

Qy      81  TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db      477 TATGAGCGTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGCGTGGCGAGGG 536
Qy      101 ArgileSerCysThrIle 106
Db      537 AGAATAAGCTGTACCATC 554

RESULT 12
LOCUS   AL706288
DEFINITION 751 bp mRNA linear EST 04-SEP-2003
ACCESSION DKFZp686P033_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
VERSION  AL706288
KEYWORDS EST.
SOURCE  Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 751)
AUTHORS  Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE  EST (Duesterhoeft, et al.)
JOURNAL  Unpublished (1999)
COMMENT  Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp686P033) is available at the R2PD in Berlin.
Please contact the R2PD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES             source
    source
    1..751
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="DKFZp686P033"
        /dev_stage="adult"
        /lab_host="DH10B"
        /clone_lib="686 (synonym: hlcc3)"
        /note="Vector: pTriplEx2; Site_1: SfIIA; Site_2: SfiIB;
        cDNA-collection"

ORIGIN
Alignment Scores:
Pred. No.:      5 22e-61      Length:      751
Score:          596.00      Matches:      105
Percent Similarity: 99.06%      Conservative: 0
Best Local Similarity: 99.06%      Mismatches: 1
Query Match:     99.33%      Indels:      0
DB:              1          Gaps:      0

US-09-940-235-4_COPY_1_106 (1-106) x AL706288 (1-751)
Qy      1  GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db      361 CAGGCTCAGCAAAATGGTTGAGCCCGGAGTCCCGGCTGCTGCTAGTCAAAAGCAAGCCCGT 420
Qy      21  CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db      421 TGTATGACATATGGAACACATATATCAGATAAAATCAACAGTGGGAGCGGACCTACCTAGGC 480
Qy      41  AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
Db      481 AATGCGTTGGTTGTACTGTTATGGAGAACCGAGGTTTAACTGCGAGAGTAAACCT 540
Qy      61  GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80

```

```

Db      541 GAAGCTGAAGAGACTTGTCTTACCAAGTACATCGGGAACACTTACCGAGTGGGTGACACT 600
Qy      81  TyrGluArgProLysAspSerMetIleTrpAspCysThrCysAlaGlyArgGly 100
Db      601 TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGGG 660
Qy      101 ArgIleSerCysThrIle 106
Db      661 AGAATAAGCTGTACCATC 678

RESULT 13
LOCUS   BX473407
DEFINITION DKFZp686G13162_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
VERSION   BX473407
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 756)
AUTHORS   Ottenwaelder B., Obermaier B., Deutschenbaur, S., Mewes, H.W.,
           Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
           EST (Ottenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,
           et al.)
TITLE     Unpublished (2003)
JOURNAL   Contact: MIPS
COMMENT   MIPS
           Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
           This is the 5' sequence of the clone insert
           Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
           Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
           sequenced by MediGenomix (Martinsried/Germany) within the cDNA
           sequencing consortium of the German Genome Project. No sl sequence
           available.
           This clone (DKFZp686G13162) is available at the RZPD in Berlin.
           Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
           Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
         source
         1..756
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="DKFZp686G13162"
         /dev_stage="adult"
         /lab_host="DH10B"
         /clone_lib="686 (synonym: hlcc3)"
         /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
         cDNA-collection"

ORIGIN
Alignment Scores:
Pred. No.: 5,266-61 Length: 756
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x BX473407 (1-756)

Qy      1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db      361 CAGGCTCAGCAATGGTTCAGCCCCAGTCCCGGTGGCTGTCTAGTCAAGCAAGCCCGGT 420
Qy      21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db      421 TGTATTGACAAATGGAAACACATATCAATAAATCAACAGTGGGAGCGGACCTACCTAGGC 480
Qy      41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60

```

```

Db      481 AATCGCTGGTGTGTACTTGTATGAGGAAGCCGAGGTTTAACTTCGCGAGACTAAACCT 540
Qy      61 GluAlaGlnGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db      541 GAAGCTGAAGAGACTTGTCTTACCAAGTACATCGGGAACACTTACCGAGTGGGTGACACT 600
Qy      81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db      601 TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGGAGGG 660
Qy      101 ArgIleSerCysThrIle 106
Db      661 AGAATAAGCTGTACCATC 678

RESULT 14
LOCUS   BX473413
DEFINITION DKFZp686I115162_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
VERSION   BX473413
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 765)
AUTHORS   Ottenwaelder B., Obermaier B., Deutschenbaur, S., Mewes, H.W.,
           Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
           EST (Ottenwaelder, B., Obermaier, B., Deutschenbaur, S., Mewes, H.W.,
           et al.)
TITLE     Unpublished (2003)
JOURNAL   Contact: MIPS
COMMENT   MIPS
           Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
           This is the 5' sequence of the clone insert
           Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
           Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
           sequenced by MediGenomix (Martinsried/Germany) within the cDNA
           sequencing consortium of the German Genome Project. No sl sequence
           available.
           This clone (DKFZp686I115162) is available at the RZPD in Berlin.
           Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
           Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
         source
         1..765
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="DKFZp686I115162"
         /dev_stage="adult"
         /lab_host="DH10B"
         /clone_lib="686 (synonym: hlcc3)"
         /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
         cDNA-collection"

ORIGIN
Alignment Scores:
Pred. No.: 5,346-61 Length: 765
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x BX473413 (1-765)

Qy      1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db      361 CAGGCTCAGCAATGGTTCAGCCCCAGTCCCGGTGGCTGTCTAGTCAAGCAAGCCCGGT 420
Qy      21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40

```

Db 421 TGTATGACAATGGAAACACATATCAGATAAAATCAACAGTGGAGCGGACCTACCTAGGC 480
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 481 AATGCGTGGTTGTACTTGTATATGAGAGAGCCGAGGTTTAACTTCGAGAGATAAACC 540
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 541 GAAAGCTGAAGAGACTTGTCTTGCACAGTACACTGGGAACACATTACCGAGTGGGTGACACT 600
Qy 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysIleGlyAlaGlyArgGly 100
Db 601 TATGAGCGTCCFAAGAGACTCCATGATCTGGGACTGTACCTGCATCGCGGCTGGGCGAGGG 660
Qy 101 ArgIleSerCysThrIle 106
Db 661 AGAATAAGCTGTACCATC 678

RESULT 15
BX398838 911 bp mRNA linear EST 29-APR-2004
LOCUS BX398838 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1065Y122 5-PRIME, mRNA sequence.
ACCESSION BX398838
VERSION BX398838.2 GI:46878456
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 911)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30625651.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1413.r

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0D1065BE11QF1&c=1413.r>.

FEATURES

Location/Qualifiers
1..911
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1065Y122"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
Pred. No.: 6,68e-61 Length: 911
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x BX398838 (1-911)

Qy 1 GluAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20

Db 325 CAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCCGGTGGCTGTCACTCAAAGCAAGCCCGGT 384
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 385 TGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGC 444
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 445 AATGCGTGGTTGTACTTGTATATGAGAGAGCCGAGGTTTAACTTCGAGAGATAAACC 504
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 505 GAAAGCTGAAGAGACTTGTCTTGCACAGTACACTGGGAACACATTACCGAGTGGGTGACACT 564
Qy 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysIleGlyAlaGlyArgGly 100
Db 565 TATGAGCGTCCFAAGAGACTCCATGATCTGGGACTGTACCTGCATCGCGGCTGGGCGAGGG 624
Qy 101 ArgIleSerCysThrIle 106
Db 625 AGAATAAGCTGTACCATC 642

Search completed: November 6, 2004, 04:55:32
Job time : 1300.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2004, 00:02:46 ; Search time 34.8493 Seconds
(without alignments)
2161.983 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600
Sequence: 1 QAQQMVQPSPVAVSQSPG.....SMIWDCTCIGAGGRISCTI 106

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame_p2n.model -runat 03112004 174039 11280/app_query.fasta_1.1045
-Q=/cgn2_1/USPTO_spool_P/US09940235/runat_03112004 174039 11280/app_query.fasta_1.1045
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLFMT=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 @Cn1_1_177 @runat 03112004 174039 11280 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	600	100.0	7679	4	US-09-220-132-38
2	600	100.0	7680	4	Sequence 38, Appl
3	600	100.0	7680	5	Sequence 1289, Ap
4	600	100.0	7705	1	PCR-US95-09819-6
5	600	100.0	7705	2	US-08-259-569-16
6	600	100.0	7705	2	US-08-826-885-16
7	600	100.0	7803	2	US-08-551-356-1
8	600	100.0	7803	5	PCR-US93-12687-1
9	596	99.3	8044	4	US-09-566-921-135
10	162.5	27.1	567	1	US-08-142-449B-5
11	160	26.7	247	4	US-09-397-787-118
12	142	23.7	828	4	US-09-404-879A-220
Sequence 220, App					

C 13	142	23.7	828	4	US-09-338-933-220	Sequence 220, App
C 14	142	23.7	828	4	US-09-215-681-220	Sequence 220, App
C 15	142	23.7	828	4	US-09-216-003A-220	Sequence 220, App
C 16	142	23.7	828	4	US-09-667-857-220	Sequence 8, Appli
C 17	141	23.5	986	1	US-07-637-250A-8	Sequence 8, Appli
C 18	141	23.5	986	1	US-08-145-061-8	Sequence 252, App
C 19	140	23.3	501	4	US-09-404-879A-252	Sequence 252, App
C 20	140	23.3	501	4	US-09-338-933-252	Sequence 252, App
C 21	140	23.3	501	4	US-09-215-681-252	Sequence 252, App
C 22	140	23.3	501	4	US-09-216-003A-252	Sequence 252, App
C 23	140	23.3	501	4	US-09-667-857-252	Sequence 252, App
C 24	136.5	22.8	440	4	US-09-397-787-321	Sequence 321, App
C 25	132	22.0	69	1	US-08-259-569-23	Sequence 23, Appl
C 26	132	22.0	69	2	US-08-826-885-23	Patent No. 5455158
C 27	127	21.2	68	6	5455158-4	Sequence 347, App
C 28	125.5	20.9	241	4	US-09-389-681-347	Sequence 347, App
C 29	125.5	20.9	241	4	US-09-620-405B-347	Sequence 347, App
C 30	125.5	20.9	241	4	US-09-433-826B-347	Sequence 347, App
C 31	125.5	20.9	241	4	US-09-604-287A-347	Sequence 347, App
C 32	125.5	20.9	241	4	US-09-834-759-347	Sequence 347, App
C 33	125.5	20.9	241	4	US-09-590-751A-347	Sequence 347, App
C 34	125	20.8	69	6	5455158-6	Patent No. 5455158
C 35	117	19.5	60	1	US-08-259-569-25	Sequence 25, Appl
C 36	117	19.5	60	2	US-08-826-885-25	Sequence 25, Appl
C 37	117	19.5	60	6	5455158-7	Patent No. 5455158
C 38	114	19.0	69	1	US-08-259-569-19	Sequence 19, Appl
C 39	114	19.0	69	2	US-08-826-885-19	Sequence 19, Appl
C 40	112	18.7	57	1	US-08-259-569-21	Sequence 21, Appl
C 41	112	18.7	57	2	US-08-826-885-21	Sequence 21, Appl
C 42	112	18.7	57	6	5455158-5	Patent No. 5455158
C 43	108	18.0	75	1	US-08-259-569-18	Sequence 18, Appl
C 44	108	18.0	75	2	US-08-826-885-18	Sequence 18, Appl
C 45	106	17.7	56	1	US-08-259-569-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-220-132-38
; Sequence 38, Application US/09220132
; Patent No. 6508607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220,132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 7679
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-38

Alignment Scores:				
Pred. No.:	3,86e-70	Length:	7679	
Score:	600.00	Matches:	106	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	4	Gaps:	0	
US-09-940-235-4_COPY_1_106 (1-106) x US-09-220-132-38 (1-7679)				
Qy	1	GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	20	
Db	14	CAGGCTCAGCAANTGGTTACGCCAGTCCCGTGGCTGTCAGTCAAGCAAGCCCGT	73	

Alignment Scores:
Pred. No.: 3,86e-70 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x PCT-US95-09819-6 (1-7680)

Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGTTTTCAGCCCGGCTGCTCAGTCAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACATGGAACACATATCAGATTAATCAACAGTGGGAGCGGCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTGTTGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTGACAAAGTACACTGGGAAACACTTACCAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 4

US-08-259-569-16
Sequence 16, Application US/08259569
Patent No. 5679320
GENERAL INFORMATION:
APPLICANT: Vogel, Tikva
APPLICANT: Levanon, Avigdor
APPLICANT: Werber, Moshe
APPLICANT: Guy, Rachel
APPLICANT: Panet, Amos
APPLICANT: Hartman, Jacob
APPLICANT: Shaked, Hadassa
TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
OPERATING SYSTEM: PC-DOS/MS-DOS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/259,569
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 703,842
FILING DATE: 21-MAY-1991
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 7705 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 5..7681
US-08-259-569-16

Alignment Scores:
Pred. No.: 3,87e-70 Length: 7705
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-08-259-569-16 (1-7705)

Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGTTTTCAGCCCGGCTGCTCAGTCAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACATGGAACACATATCAGATTAATCAACAGTGGGAGCGGCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTGTTGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTCGAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTGACAAAGTACACTGGGAAACACTTACCAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGGCGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 5
US-08-826-885-16
Sequence 16, Application US/08826885
Patent No. 5869616
GENERAL INFORMATION:
APPLICANT: Vogel, Tikva
APPLICANT: Levanon, Avigdor
APPLICANT: Werber, Moshe
APPLICANT: Guy, Rachel
APPLICANT: Panet, Amos
APPLICANT: Hartman, Jacob
APPLICANT: Shaked, Hadassa
TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
OPERATING SYSTEM: PC-DOS/MS-DOS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA

```
;
;
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,885
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/703,842
; FILING DATE: 21-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 25775-D/JPW/EAB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7705 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5..7681
; US-08-826-885-16

Alignment Scores:
Pred. No.: 3 87e-70 Length: 7705
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-08-826-885-16 (1-7705)
QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGTTTCAGCCCCAGTCCCGGTGGTGTCTCAGTCAAGCAAGCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAACACATATCAGATAAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTTGTGTTTGTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTCTTCACAAAGTACACCTGGGAACACATTACCGAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCACTCGGGGCTGGCGGAGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 6
US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
;
;
; GUY, RACHEL; PANET, AMOS
; TITLE OF INVENTION: FIBRIN BINDING DOMAIN POLYPEPTIDES AND
; USES AND METHODS OF PRODUCING SAME
; NUMBER OF SEQUENCES: 20
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,241
; FILING DATE: 04-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 526,397
; FILING DATE: 21-MAY-1990
; APPLICATION NUMBER: 345,952
; FILING DATE: 28-APR-1989
; APPLICATION NUMBER: 291,951
; FILING DATE: 29-DEC-1988
; SEQ ID NO: 2:
; LENGTH: 7705
; 5455158-2

Alignment Scores:
Pred. No.: 3 87e-70 Length: 7705
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x 5455158-2 (1-7705)
QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGTTTCAGCCCCAGTCCCGGTGGTGTCTCAGTCAAGCAAGCCGGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAAATGGAACACATATCAGATAAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTTGTGTTTGTACTTGTATGAGGAAGCCGAGGTTTAACTGCGAAAGTAACCT 193
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTCTTCACAAAGTACACCTGGGAACACATTACCGAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCACTCGGGGCTGGCGGAGG 313
QY 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 7
US-08-551-356-1
; Sequence 1, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/551,356
FILING DATE: 31-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/349,762
FILING DATE: 31-DEC-1992
APPLICATION NUMBER: US/07/998,271
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Holly, Julie A
REGISTRATION NUMBER: 33-246
REFERENCE/DOCKET NUMBER: 92-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 6..7346
US-08-551-356-1

Alignment Scores:
Pred. No.: 3.94e-70 Length: 7803
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-08-551-356-1 (1-7803)

Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 99 CAGGCTCAGCAATGGTTTCAGCCCAAGTCCCGGTGGTGTCTAGTCAAGCAAGCCGGT 158
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 40
Db 159 TGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACTACCTAGGT 218
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 219 AATGTTGGTTGTACTTGTATGGAGGAGCGGAGGTTTAACTGCCGAAAGTAAACCT 278
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 279 GAAGCTGAAGAGACTTGTCTTGCACAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 338
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 339 TATGAGCGTCTAAAGACTCCATGATCTGGGACTGATCTGGGAGCGGCTCGGGCGGAGGG 398
Qy 101 ArgIleSerCysThrIle 106
Db 399 AGAATAAGCTGTACCATC 416

RESULT 8

PCT-US93-12687-1
Sequence 1, Application PC/TUS9312687
GENERAL INFORMATION:
APPLICANT: Irani, Meher H.
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA

ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12687
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/998,271
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 92-26PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 6..7346
PCT-US93-12687-1

Alignment Scores:
Pred. No.: 3.94e-70 Length: 7803
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x PCT-US93-12687-1 (1-7803)

Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 99 CAGGCTCAGCAATGGTTTCAGCCCAAGTCCCGGTGGTGTCTAGTCAAGCAAGCCGGT 158
Qy 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 40
Db 159 TGTATGACAAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACTACCTAGGT 218
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 219 AATGTTGGTTGTACTTGTATGGAGGAGCGGAGGTTTAACTGCCGAAAGTAAACCT 278
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 279 GAAGCTGAAGAGACTTGTCTTGCACAGTACACTGGGAAACACTTACCGAGTGGGTGACACT 338
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 339 TATGAGCGTCTAAAGACTCCATGATCTGGGACTGATCTGGGAGCGGCTCGGGCGGAGGG 398
Qy 101 ArgIleSerCysThrIle 106
Db 399 AGAATAAGCTGTACCATC 416

RESULT 9

US-09-566-921-135
Sequence 135, Application US/09566921
Patent No. 668288
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne P.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.

;; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
;; FILE REFERENCE: PA-0024 US
;; CURRENT APPLICATION NUMBER: US/09/566,921
;; CURRENT FILING DATE: 2000-05-05
;; NUMBER OF SEQ ID NOS: 138
;; SOFTWARE: PERL Program
;; SEQ ID NO 135
;; LENGTH: 8044
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; OTHER INFORMATION: Incyte ID No. 6582888 427813.14
US-09-566-921-135

Alignment Scores:
Pred. No.: 1.43e-69 Length: 8044
Score: 596.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 99.33% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-09-566-921-135 (1-8044)

Qy 1 GlnAlaGlnMetValClnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 362 CAGGCTCAGCAAAATGGTTCAGCCCGAGTCCCGGGTGTGTAGTCAAAAGCAAGCCCGT 421
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 40
Db 422 TGTATGACATGGAACACATATCAGATAATCAACATGGGAGCGGACCTACCTAGGC 481
Qy 41 AsnValLeuValCysThrCysTyrGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 482 AATGGTGTGGTGTGTACTGTATGAGGAGCGAGGTTTAACTGCGAGAGTAACCT 541
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 542 GAAGCTGAAGAGACTTGCTTTGACAAAGTACATCGGAACACATTTACCGAGTGGGTGACCT 601
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 602 TATGAGCGTCTTAAGACTCCATGATCTGGGACTGTACCTGCAATGGGCTGGCGAGGG 661
Qy 101 ArgIleSerCysThrIle 106
Db 662 AGAATAAGCTGTACCATC 679

RESULT 10
US-08-142-449B-5
; Sequence 5, Application US/08142449B
; Patent No. 5668104
; GENERAL INFORMATION:
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kawano, Genji
; APPLICANT: Sudo, Tetsuo
; APPLICANT: Kojima, Katsuaki
; TITLE OF INVENTION: Physiologically Active Protein and
; TITLE OF INVENTION: Hematopoietic Stem Cell Growth Agent
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nels T. Lippert, White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/142,449B
;; FILING DATE: 24-NOV-1994
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lippert, Nels T.
;; REGISTRATION NUMBER: 25,888
;; REFERENCE/DOCKET NUMBER: 1145358-304
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (212)-819-8582
;; TELEFAX: (212) 354-8113
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 567 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA to mRNA
;; FEATURE:
;; NAME/KEY: CDS
US-08-142-449B-5

Alignment Scores:
Pred. No.: 1.04e-12 Length: 567
Score: 162.50 Matches: 32
Percent Similarity: 57.32% Conservative: 15
Best Local Similarity: 39.02% Mismatches: 28
Query Match: 27.08% Indels: 7
DB: 1 Gaps: 4

US-09-940-235-4_COPY_1_106 (1-106) x US-08-142-449B-5 (1-567)

Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThr---TyrLeu 39
Db 157 TGGCATGACAATGGTGTGAACACTACAAATGGAGAGAGTGGACCGTCAGGAGAGAAAT 216
Qy 40 GlyAsnValLeuValCysThrCysTyrGlyGlySerArgGly---PheAsnCysGluSer 58
Db 217 GGCAGATGATGAGTGCACATGTCTTGGAGCGGAAAGAGAAATCAAGTGTGACCT 276
Qy 59 LysProGluAlaGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGly 78
Db 277 CAT-----GAGCAACGCTGTACGAT-----GATGGGAAGACATACCACGTAGGA 321
Qy 79 AspThrTyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGly 98
Db 322 GAACAGTGGCAGGAAGGAATATCTCGTGGCAATTTGCTCTGCACATGCTTTGGAGGCCAG 381
Qy 99 ArgGly 100
Db 382 CGGGGC 387

RESULT 11
US-09-397-787-118
; Sequence 118, Application US/093977787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 118
; LENGTH: 247
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-118

```
Alignment Scores:
Pred. No.: 7,07e-13 Length: 247
Score: 160.00 Matches: 26
Percent Similarity: 68.42% Conservatives: 13
Best Local Similarity: 45.61% Mismatches: 16
Query Match: 26.67% Indels: 2
DB: 4 Gaps: 1

US-09-940-235-4_COPY_1_106 (1-106) x US-09-397-787-118 (1-247)

Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 40
Db 25 TGCTATGACGATGGGAGACCTACCATGTAGGAGAACAGTGGCAGAGGAATATCTCGG 84
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySer----- 58
Db 85 GCCATTTGCTCTCGACGATGTTTCGGAGCCGCGGGCTGGCGCTGTGACAACTGCCGT 144
Qy 59 LysProGluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyr 75
Db 145 AGACCTGGGGCTGCTGACCCAGTCCCGATGGCACCACCGGCCACACCTAC 195

RESULT 12
US-09-404-879A-220/c
; Sequence 220, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(828)
; OTHER INFORMATION: n = A,T,C or G
US-09-404-879A-220

Alignment Scores:
Pred. No.: 1.03e-09 Length: 828
Score: 142.00 Matches: 22
Percent Similarity: 64.71% Conservatives: 11
Best Local Similarity: 43.14% Mismatches: 12
Query Match: 23.67% Indels: 6
DB: 4 Gaps: 1

US-09-940-235-4_COPY_1_106 (1-106) x US-09-404-879A-220 (1-828)

Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 40
Db 180 TGTTCAGTATGATGGGAGACATACCATCGTAGGAGAACAGTGGCAGAGGAATATCTCGT 121
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySer----- 58
Db 120 GCCATTTGCTCTCGACATGCTTTGGAGCCGCGGGCTGGCGCTGTGACAACTGCCGC 61
Qy 59 -----LysProGluAlaGluGluThr 65
Db 60 AGACCTGGGGGTGACCCAGTCCCGAAGGCACT 28

RESULT 13
US-09-933-933-220/c
; Sequence 220, Application US/09338933
```

```
Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(828)
; OTHER INFORMATION: n = A,T,C or G
US-09-338-933-220

Alignment Scores:
Pred. No.: 1.03e-09 Length: 828
Score: 142.00 Matches: 22
Percent Similarity: 64.71% Conservatives: 11
Best Local Similarity: 43.14% Mismatches: 12
Query Match: 23.67% Indels: 6
DB: 4 Gaps: 1

US-09-940-235-4_COPY_1_106 (1-106) x US-09-338-933-220 (1-828)

Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly 40
Db 180 TGTTCAGTATGATGGGAGACATACCATCGTAGGAGAACAGTGGCAGAGGAATATCTCGT 121
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySer----- 58
Db 120 GCCATTTGCTCTCGACATGCTTTGGAGCCGCGGGCTGGCGCTGTGACAACTGCCGC 61
Qy 59 -----LysProGluAlaGluGluThr 65
Db 60 AGACCTGGGGGTGACCCAGTCCCGAAGGCACT 28

RESULT 14
US-09-215-681-220/c
; Sequence 220, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Frudakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF OVARIAN CANCER
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 220
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(828)
; OTHER INFORMATION: n = A,T,C or G
US-09-215-681-220

Alignment Scores:
Pred. No.: 1.03e-09 Length: 828
Score: 142.00 Matches: 22
Percent Similarity: 64.71% Conservatives: 11
Best Local Similarity: 43.14% Mismatches: 12
```

Alignment Scores:		
Pred. No.:	1.03e-09	Length: 828
Score:	142.00	Matches: 22

Percent Similarity: 64.71% Conservative: 11
 Best Local Similarity: 43.14% Mismatches: 12
 Query Match: 23.67% Indels: 6
 DB: 4 Gaps: 1

US-09-940-235-4_COPY_1_106 (1-106) x US-09-216-003A-220 (1-828)

Qy	21	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly	40
Db	180	TGTTACGATGATGGGAAGACATACCACGTAGGAGAACAGTGGCAGAAGGAATATCTCGGT	121
Qy	41	AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSer-----	58
Db	120	GCCATTTCCTCTGCACATGCTTTGGAGGCCAGCGGGGCTGGCGCTGTGACAACTGCCGC	61
Qy	59	-----LysProGluAlaGluGluThr	65
Db	60	AGACCTGGGGGTGAACCCAGTCCCGAAGGCACT	28

Search completed: November 6, 2004, 04:59:09
 Job time : 41.8493 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2004, 02:55:27 ; Search time 182.233 Seconds
(without alignments)
3131.797 Million cell updates/sec

Title: US-09-940-235-4_COPY_1_106

Perfect score: 600

Sequence: 1 QAQMVPQSPVAVSQSRPG.....SMWDCTCTGAGRISCTI 106

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spc/p/US09940235/runat_03112004_174041_11368/app_query.fasta_1.1045
-DB=Published Applications NA -QMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09940235@cgn_1_1_1034@runat_03112004_174041_11368
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCk=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB ID	Description
1	600	100.0	777	10	US-09-940-235-3
2	600	100.0	1661	10	US-09-940-235-10
3	600	100.0	6988	16	US-10-236-392-1
4	600	100.0	7680	9	US-09-964-824A-574
5	600	100.0	7680	14	US-10-171-311-63
6	600	100.0	7680	15	US-10-236-031B-69
7	600	100.0	7680	15	US-10-374-979-75
8	600	100.0	7680	16	US-10-182-936A-75
9	600	100.0	7680	16	US-10-641-643-1289
10	600	100.0	7680	17	US-10-717-597-222
11	600	100.0	7680	18	US-10-788-792-79
12	600	100.0	7705	16	US-10-447-161-4
13	596	99.3	867	9	US-09-925-302-105
14	596	99.3	867	10	US-09-925-302-105
15	596	99.3	2127	15	US-10-210-120-49
16	596	99.3	2443	17	US-10-741-601-70
17	596	99.3	2488	17	US-10-741-601-75
18	596	99.3	4295	15	US-10-144-194A-51
19	596	99.3	6510	17	US-10-741-601-72
20	596	99.3	7361	16	US-10-236-392-3
21	596	99.3	7795	15	US-10-084-817-2
22	596	99.3	7823	17	US-10-741-601-77
23	596	99.3	7848	17	US-10-741-601-78
24	596	99.3	7867	13	US-10-098-841-6
25	596	99.3	7935	17	US-10-741-601-74
26	596	99.3	7959	17	US-10-741-601-81
27	596	99.3	8013	17	US-10-741-601-71
28	596	99.3	8027	16	US-10-447-161-8
29	596	99.3	8027	17	US-10-734-564-27
30	596	99.3	8044	15	US-10-240-965-121
31	596	99.3	8062	13	US-10-098-841-5
32	596	99.3	8137	13	US-10-098-841-8
33	596	99.3	8155	17	US-10-741-601-79
34	596	99.3	8226	17	US-10-741-601-69
35	596	99.3	8230	13	US-10-098-841-7
36	596	99.3	8278	17	US-10-741-601-82
37	596	99.3	8332	17	US-10-741-601-73
38	596	99.3	8371	17	US-10-741-601-76
39	574	95.7	8329	9	US-09-917-800A-1731
40	574	95.7	8329	16	US-10-191-803-113
41	317	52.8	407	9	US-09-728-445-799
42	271.5	45.2	87467	17	US-10-741-601-5634
43	253	42.2	201	17	US-10-741-601-1761
44	253	42.2	201	17	US-10-741-601-1764
45	253	42.2	201	17	US-10-741-601-1779

ALIGNMENTS

RESULT 1
US-09-940-235-3
; Sequence 3, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammarra
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23

; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3

; LENGTH: 777
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(777)
US-09-940-235-3

Alignment Scores:
Pred. No.: 6,178-76 Length: 777
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-09-940-235-3 (1-777)

QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 1 CAGGCTCAGCAATGGTTCAGCCCGAGTCCCGGTGGTGTCTCAGTCAAGCAAGCCCGGT 60
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 61 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 120
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
DB 121 AATGTTGGTGTGTACTTGTATGGAGGAGCGGAGGTTTAACTGCCAAAGTAAACCT 180
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 181 GAAGCTGAAGAGACTTGTTCAGCAAGTACATCTGGGAACACTTACCGAGTGGGTGACCT 240
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 241 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGGGCTGGGCGAGGG 300
QY 101 ArgIleSerCysThrIle 106
DB 301 AGAATAAGCTGTACCATC 318

RESULT 2

US-09-940-235-10
; Sequence 10, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir

; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10

; LENGTH: 1661
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hybrid cassette
US-09-940-235-10

Alignment Scores:
Pred. No.: 1,748-75 Length: 1661
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-09-940-235-10 (1-1661)

QY 1 GlnAlaGlnGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 1341 CAGGCGCAGCAATGGTTCAGCCCGAGTCCCGGTGGTGTCTCAGTCAAGCAAGCCCGGT 1460
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 1401 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 1460
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
DB 1461 AATGTTGGTGTGTACTTGTATGGAGGAGCGGAGGTTTAACTGCCAAAGTAAACCT 1520
QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 1521 GAAGCTGAAGAGACTTGTTCAGCAAGTACATCTGGGAACACTTACCGAGTGGGTGACCT 1580
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 1581 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTACCTGCGGGCTGGGCGAGGG 1640
QY 101 ArgIleSerCysThrIle 106
DB 1641 AGAATAAGCTGTACCATC 1658

RESULT 3

US-10-236-392-1
; Sequence 1, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRocheville, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E

APPLICANT: Shenoy, Suresh
APPLICANT: Shimkets, Richard A
APPLICANT: Smithson, Glenda
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-442A
CURRENT APPLICATION NUMBER: US/10/236,392
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: US09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US60/390,155
PRIOR FILING DATE: 2002-06-19
PRIOR APPLICATION NUMBER: US09/635,949
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US60/318,765
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: US60/357,303
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US60/367,753
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: US60/369,479
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: US09/659,634
PRIOR FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: US60/318,120
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US60/318,130
PRIOR FILING DATE: 2001-09-07
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 794
SOFTWARE: Custom
SEQ ID NO 1
LENGTH: 6988
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (26)..(6986)
US-10-236-392-1

Alignment Scores:
Pred. No.: 1,24e-74 Length: 6988
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-236-392-1 (1-6988)
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAGTCAAAAGCAAGCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAATGGAAAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTTGGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGCCTTTCAGCAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGCGGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 4
US-09-964-824A-574
Sequence 574, Application US/09964824A
Patent No. US20020102531A1
GENERAL INFORMATION:
APPLICANT: Horrigan, Stephen
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
FILE REFERENCE: 689290-73
CURRENT APPLICATION NUMBER: US/09/964,824A
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/60/236,033
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,032
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US/60/236,028
PRIOR FILING DATE: 2000-09-28
NUMBER OF SEQ ID NOS: 583
SOFTWARE: PatentIn version 3.0
SEQ ID NO 574
LENGTH: 7680
TYPE: DNA
ORGANISM: Homo sapiens
US-09-964-824A-574
Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-940-235-4_COPY_1_106 (1-106) x US-09-964-824A-574 (1-7680)
Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAAAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAGTCAAAAGCAAGCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnInTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAATGGAAAAACACTATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro 60
Db 134 AATGTTGGTTTGTACTTGTATGAGGAGCGGAGGTTTAACTGCGAAAGTAAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGCCTTTCAGCAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTTPAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCCCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCTGGCGGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331
RESULT 5
US-10-171-311-63
Sequence 63, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Glatt, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoerh, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

```
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: RastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-63

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-171-311-63 (1-7680)
QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAATGGTTTCAGCCCGTCCCGTGGCTGCTCAGTCAAGCAAGCCCGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATGACAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
DB 134 AATGTTGGTTTGTACTTGTATGAGGAGGAGCGAGGTTTAACTGCGAAAGTAAACCT 193
QY 61 GluAlaGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCGCTGGCGGAGG 313
QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331

RESULT 6
US-10-236-031B-69
; Sequence 69, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: RastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-63

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-171-311-63 (1-7680)
QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAATGGTTTCAGCCCGTCCCGTGGCTGCTCAGTCAAGCAAGCCCGT 73
QY 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATGACAATGGAAACACATATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGT 133
QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlySerLysPro 60
DB 134 AATGTTGGTTTGTACTTGTATGAGGAGGAGCGAGGTTTAACTGCGAAAGTAAACCT 193
QY 61 GluAlaGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGACAAGTACACTGGGAACACTTACCGAGTGGGTGACACT 253
QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTCTAAAGACTCCATGATCTGGGACTGTACCTGCGGCGCTGGCGGAGG 313
QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331

RESULT 7
US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-75

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
```

US-09-940-235-4_COPY_1_106 (1-106) x US-10-374-979-75 (1-7680)

Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGGTTGAGCCCGAGTCCCGGGTGTCTAGTCAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
Db 74 TGTATGACAATGGAAAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGlnSerLysPro 60
Db 134 AATGTGTGGTTGTACTTGTATGGAGAGCCGAGGTTTAACTCGCAAGTAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTGGACAAAGTACACTGGGAACACTTACCGAGTGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAGAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGCGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 8
US-10-182-936A-75
; Sequence 75, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Sheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-75

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-182-936A-75 (1-7680)

Qy 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
Db 14 CAGGCTCAGCAATGGTTGAGCCCGAGTCCCGGGTGTCTAGTCAAGCAAGCCCGGT 73
Qy 21 CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40

Db 74 TGTATGACAATGGAAAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT 133
Qy 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGlnSerLysPro 60
Db 134 AATGTGTGGTTGTACTTGTATGGAGAGCCGAGGTTTAACTCGCAAGTAACCT 193
Qy 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
Db 194 GAAGCTGAAGAGACTTGTCTTGGACAAAGTACACTGGGAACACTTACCGAGTGGTGACACT 253
Qy 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
Db 254 TATGAGCGTCTTAAGAGACTCCATGATCTGGGACTGTACTGTCATCGGGCTGGCGAGGG 313
Qy 101 ArgIleSerCysThrIle 106
Db 314 AGAATAAGCTGTACCATC 331

RESULT 9
US-10-641-643-1289
; Sequence 1289, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/641,643
; FILING DATE: 14-Aug-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g31396
; SEQUENCE DESCRIPTION: SEQ ID NO: 1289 :

US-10-641-643-1289

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-641-643-1289 (1-7680)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAAAATGGTTTTCAGCCCGAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCGGT 73

QY -21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATGATCAATGGAAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133

QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlyPheAsnGlySerLysPro 60
DB 134 AATGTTGGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 193

QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGCACAGTACACTTGGGAACACATTACCGAGTGGGTGACACT 253

QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGCGAGGG 313

QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331

RESULT 10
US-10-717-597-222
; Sequence 222, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dornek, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM101080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US 60/459,782
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: US 60/427,982
; PRIOR FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 222
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-717-597-222

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-717-597-222 (1-7680)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAAAATGGTTTTCAGCCCGAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCGGT 73

QY 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40

DB 74 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133

QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlyPheAsnGlySerLysPro 60
DB 134 AATGTTGGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 193

QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGCACAGTACACTTGGGAACACATTACCGAGTGGGTGACACT 253

QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGCGAGGG 313

QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331

RESULT 11
US-10-788-792-79
; Sequence 79, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 79
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-788-792-79

Alignment Scores:
Pred. No.: 1,41e-74 Length: 7680
Score: 600.00 Matches: 106
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-788-792-79 (1-7680)

QY 1 GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly 20
DB 14 CAGGCTCAGCAAAATGGTTTTCAGCCCGAGTCCCGGTGGCTGTGCTCAGTCAAGCAAGCCCGGT 73

QY 21 CysTyrAspAsnGlyHisTyrGlnIleAsnGlnGlnTrpGluArgThrTyrLeuGly 40
DB 74 TGTATGACAAATGGAAACACATATCAGATAAATCAACAGTGGAGCGGACCTACCTAGGT 133

QY 41 AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnGlyPheAsnGlySerLysPro 60
DB 134 AATGTTGGTTTGTACTTGTATGGAGGAGCCGAGGTTTAACTGCGAAAGTAAACCT 193

QY 61 GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr 80
DB 194 GAAGCTGAAGAGACTTGTCTTGCACAGTACACTTGGGAACACATTACCGAGTGGGTGACACT 253

QY 81 TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly 100
DB 254 TATGAGCGTCTTAAAGACTCCATGATCTGGGACTGTGACATCGGGCTGGGCGAGGG 313

QY 101 ArgIleSerCysThrIle 106
DB 314 AGAATAAGCTGTACCATC 331

RESULT 12

US-10-447-161-4
; Sequence 4, Application US/10447161
; Publication No. US20040023314A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Rong-fu
; TITLE OF INVENTION: Mutant Fibronectin and Tumor Metastasis
; FILE REFERENCE: HO-P02484US1
; CURRENT APPLICATION NUMBER: US/10/447,161
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/383,530
; PRIOR FILING DATE: 2002-05-28
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 7705
; TYPE: DNA
; ORGANISM: Human
US-10-447-161-4

Alignment Scores:

Pred. No.:	1,41e-74	Length:	7705
Score:	600.00	Matches:	106
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	16	Gaps:	0

US-09-940-235-4_COPY_1_106 (1-106) x US-10-447-161-4 (1-7705)

QY	1	GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	20
DB	14	CAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT	73
QY	21	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly	40
DB	74	TGTTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGT	133
QY	41	AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	60
DB	134	AATGCGTGGTTGTACTTGTATGGAGGAGCGGAGGTTTAACTCGCAAGTAAACCT	193
QY	61	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	80
DB	194	GAAGCTGAAGAGACTTGTTCAGCAAGTACACTGGGAACACTTACCAGTGGGTGACACT	253
QY	81	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly	100
DB	254	TATGAGCGTCTTAAGAGCTCCATGATCTGGGACTGTACCTGCATCGGGGCTGGCGAGGG	313
QY	101	ArgIleSerCysThrIle	106
DB	314	AGAATAAGCTGTACCATC	331

RESULT 13

US-09-925-302-105
; Sequence 105, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:

Pred. No.:	2,68e-75	Length:	867
Score:	596.00	Matches:	105
Percent Similarity:	99.06%	Conservative:	0
Best Local Similarity:	99.06%	Mismatches:	1
Query Match:	99.33%	Indels:	0
DB:	10	Gaps:	0

US-09-940-235-4_COPY_1_106 (1-106) x US-09-925-302-105 (1-867)

QY	1	GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	20
DB	355	CAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT	414

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:

Pred. No.:	2,68e-75	Length:	867
Score:	596.00	Matches:	105
Percent Similarity:	99.06%	Conservative:	0
Best Local Similarity:	99.06%	Mismatches:	1
Query Match:	99.33%	Indels:	0
DB:	9	Gaps:	0

US-09-940-235-4_COPY_1_106 (1-106) x US-09-925-302-105 (1-867)

QY	1	GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	20
DB	355	CAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT	414
QY	21	CysTyrAspAsnGlyLysHisTyrGlnIleAsnGlnTrpGluArgThrTyrLeuGly	40
DB	415	TGTTATGACAATGGAAACACTATCAGATAAATCAACAGTGGGAGCGGACCTACCTAGGC	474
QY	41	AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	60
DB	475	AATGCGTGGTTGTACTTGTATGGAGGAGCGGAGGTTTAACTCGGAGAGTAAACCT	534
QY	61	GluAlaGluGluThrCysPheAspLysTyrThrGlyAsnThrTyrArgValGlyAspThr	80
DB	535	GAAGCTGAAGAGACTTGTTCAGCAAGTACACTGGGAACACTTACCAGTGGGTGACACT	594
QY	81	TyrGluArgProLysAspSerMetIleTrpAspCysThrCysIleGlyAlaGlyArgGly	100
DB	595	TATGAGCGTCTTAAGAGCTCCATGATCTGGGACTGTACCTGCATYGGGGCTGGCGAGGG	654
QY	101	ArgIleSerCysThrIle	106
DB	655	AGAATAAGCTGTACCATC	672

RESULT 14

US-09-925-302-105
; Sequence 105, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-302-105

Alignment Scores:

Pred. No.:	2,68e-75	Length:	867
Score:	596.00	Matches:	105
Percent Similarity:	99.06%	Conservative:	0
Best Local Similarity:	99.06%	Mismatches:	1
Query Match:	99.33%	Indels:	0
DB:	10	Gaps:	0

US-09-940-235-4_COPY_1_106 (1-106) x US-09-925-302-105 (1-867)

QY	1	GlnAlaGlnMetValGlnProGlnSerProValAlaValSerGlnSerLysProGly	20
DB	355	CAGGCTCAGCAATGGTTTCAGCCCCAGTCCCGGTGGCTGTCAGTCAAGCAAGCCCGGT	414

Qy		21	CysTyrAspAsnGlyLysHisTyrClnlleAsnGlnGlnTrpGluArgThrTyrLeuGly	40
Dd		415	TGTATGAACAATGGAAAAACACTATCAGATAAATCAACAGTGGAGCGGACCCTACCTAGGC	474
Qy		41	AsnValLeuValCysThrCysTyrGlyGlySerArgGlyPheAsnCysGluSerLysPro	60
Dd		475	AATCGGTGGTTGTGTACTTGTTATCGAGAGAAGCCGAGGTTTTAACTTCGCAGAGTAACCT	534
Qy		61	GluAlaGluGluThrCysPheAspIysVtYrThrGlyAsnThrTyArgValGlyAspThr	80
Dd		535	GAAGCTGAAGAGACTTGTTGTGACAAAGTACACTCGGAGACACTTACCGAGTGGGTGACACT	594
Qy		81	TyrGluArgProLysAspSerMetileTrpAspCysThrCysiledylalaGlyArgGly	100
Dd		595	TATGAGCGTCTCAAAGACTCATGATCTGGGACTGTACCTGCTATYGGGGCTGGCGAGGG	654
Qy		101	ArgIleSerCysThrIle	106
Dd		655	AGAATAAGCTGTACCATC	672

RESULT 15

```

US-10-120-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-49

```

Alignment Scores:
Pred. No.: 9.13e-75
Score: 596.00
Percent Similarity: 99.06%
Best Local Similarity: 99.06% Matches: 105
Query Match: 99.33% Mismatches: 0
DB: 15 Indels: 0
Gaps: 0
Length: 2127

US-09-940-235-4 COPY 1 106 (1-106) x US-10-210-120-49 (1-2127)

Qy	1	Gln	Ala	Gln	Gln	Met	Val	Gln	Pro	Gln	Ser	Pro	Val	Ala	Val	Ser	Cln	Ser	lys	Pro	Gly	20	
Db	150	CAG	GCT	CAG	CAA	TGG	TTT	CAG	CCC	CAG	TCCC	CGG	TGG	CTG	TG	CAG	TCAA	AG	CAAG	CCCC	GGT	209	
Qy	21	Cys	Tyr	Asp	Asn	Gly	Val	Ser	His	Tyr	Gln	Ile	Asn	Gln	Gln	Tyr	Glu	Arg	Thr	Tyr	Ileu	Gly	40
Db	210	TGT	TAT	GAC	AA	TG	AAA	ACA	CAC	TAT	CAG	ATA	AA	TCA	CA	AGT	GCG	GAC	TAC	CTA	GAC	269	
Qy	41	Asn	Val	Ileu	Val	Cys	Thr	Cys	Tyr	Gly	Gly	Ser	Arg	Gly	Phe	Asn	Cys	Glu	Ser	lys	Pro	60	
Db	270	AAT	CGG	TGG	TTG	TG	ACT	TGT	TAT	GCG	AGA	CCG	CAG	GTT	TTT	TAA	CTG	CAG	AGT	AA	CACT	329	
Qy	61	Glu	Ala	Glu	Glu	Thr	Cys	Phe	Asp	lys	Tyr	Thr	Gly	Asn	Thr	Tyr	Arg	Val	Gly	Asp	Thr	80	
Db	330	GA	AGCT	GA	AG	A	GAC	ACT	TG	CTT	TG	ACA	CA	AGT	ACA	CTG	GGA	CAC	TAC	CG	AGT	389	
Qy	81	Tyr	Glu	Arg	Pro	Lys	Asp	Ser	Met	Ile	Thr	Asp	Cys	Thr	Cys	Ile	Gly	Ala	Gly	Arg	Gly	100	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:33:59 ; Search time 53.4931 seconds
(without alignments)
737.669 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIARKCFDHAAGTSYVVGET.....ERHTSVQTTSGSGPFTDVR 110

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	627	100.0	259	3 AAY90281	Aay90281 Human fib
2	627	100.0	463	6 ABR58303	Abr58303 BCU0770 p
3	627	100.0	660	2 AAY28901	Aay28901 Human mig
4	627	100.0	720	2 AAY28914	Aay28914 Fibronect
5	627	100.0	1179	8 ADP75952	Adp75952 Human min
6	627	100.0	1359	8 ADP75957	Adp75957 Human leu
7	627	100.0	2220	6 ABO01289	Abo01289 Human pro
8	627	100.0	2220	8 ADM95950	Adm95950 Human MOV
9	627	100.0	2265	4 AAM38647	Aam38647 Human pol
10	627	100.0	2266	6 ABR40124	Abr40124 Human cel
11	627	100.0	2320	6 AAE37107	Aae37107 Human fib
12	627	100.0	2320	6 ABR42588	Abr42588 Human fib
13	627	100.0	2320	6 ABO01288	Abo01288 Human pro
14	627	100.0	2324	2 AAR92778	Aar92778 Human fib
15	627	100.0	2324	5 AAU74674	Aau74674 Human fib
16	627	100.0	2324	5 AAE23651	Aae23651 Human pro
17	627	100.0	2328	4 AAG68182	Aag68182 Fibronect
18	627	100.0	2328	6 ABU07486	Abu07486 Protein d
19	627	100.0	2328	6 ABR41106	Abr41106 Human fib
20	627	100.0	2328	6 ABR92078	Abr92078 Human cer
21	627	100.0	2328	7 ADB70378	Adb70378 Fibronect
22	627	100.0	2328	7 ADS98726	Ads98726 Human fib
23	627	100.0	2328	7 ADE82522	Ade82522 Human pro
24	627	100.0	2328	8 ADJ37157	Adj37157 Human mal
25	627	100.0	2330	4 AAM38646	Aam38646 Human pol

ALIGNMENTS

RESULT 1

AAY90281

ID AAY90281 standard; protein; 259 AA.

AC AAY90281;

DT 13-OCT-2000 (first entry)

DE Human fibronectin protein sequence fragment.

XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
KW plasminogen; human; fibronectin; thrombolytic therapy;
KW cardiovascular disorder; fibronectin.

OS Homo sapiens.

PN EP1024192-A2.

XX 02-AUG-2000.

XX 23-DEC-1999; 99EP-00310541.

XX 24-DEC-1998; 98IN-DE003825.

PR (COUL) CSIR COUNCIL SCI IND RES.

PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
Yadav M;

DR WPI; 2000-516032/47.

DR N-PSDB; AAA37632.

XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
thrombolytic therapy comprises a streptokinase fused with fibrin binding
domains of human fibronectin.

PS Example 3; Fig 6; 58pp; English.

CC This sequence represents a human fibronectin fragment, containing fibrin
binding domains. The invention relates to a hybrid plasminogen activator
(PA) comprises a polypeptide fusion between streptokinase (SK), which are
capable of plasminogen (PG) activation, and fibrin binding regions of
human fibronectin, which are from fibrin binding domains (PBD) 4 and 5 or
1 and 2. The hybrid PA possesses the ability to bind with fibrin
independently and also characteristically retains a PG activation ability
which becomes evident only after a pronounced duration, or lag, after
exposure of the PA to a suitable animal or human PG. The hybrid
streptokinase-fibrin binding domain polypeptides are useful in

Aam38649 Human pol
Abr58335 NM_00202
Adp65196 Human fib
Adg89560 Human fib
Adl92160 Fibronect
Ado55175 Protein #
Adq26085 Fibronect
Aaw63171 Amino aci
Aam38648 Human pol
Aao17353 Human fib
Abr81866 Human fib
Add18770 Human dis
Ade63324 Human-Pro
Ado55174 Protein #
Aar60021 Fibrinoge
Aab50377 Human fib
Aam40434 Human pol
Aam40435 Human pol
Aam40433 Human pol
Aam40432 Human pol

26 627 100.0 2355 4 AAM38649
27 627 100.0 2355 6 ABR58335
28 627 100.0 2355 7 ADP65196
29 627 100.0 2355 8 ADG89560
30 627 100.0 2355 8 ADL92160
31 627 100.0 2355 8 ADO55175
32 627 100.0 2355 8 ADQ26085
33 627 100.0 2386 2 AAW63171
34 627 100.0 2386 4 AAM38648
35 627 100.0 2386 5 AAO17353
36 627 100.0 2386 6 ABR81866
37 627 100.0 2386 7 ADD18770
38 627 100.0 2386 7 ADE63324
39 627 100.0 2386 8 ADO55174
40 627 100.0 2446 2 AAR60021
41 627 100.0 2446 3 AAB50377
42 627 100.0 2447 4 AAM40434
43 627 100.0 2447 4 AAM40435
44 627 100.0 2447 4 AAM40433
45 627 100.0 2447 4 AAM40432

QY 61 RIGDTSKDKNRGNLLQICITGNGRGEWKRCERTSVQTTSSGSGPFTDVR 110
|||||
Db 259 RIGDTSKDKNRGNLLQICITGNGRGEWKRCERTSVQTTSSGSGPFTDVR 308

RESULT 4

AAV28914
ID AAV28914 standard; protein; 720 AA.
AC AAV28914;
XX
DT 21-SEP-1999 (first entry)
XX
DE Fibronectin protein sequence.
XX
KW Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.
XX
XX Homo sapiens.
XX OS
XX PN WO9931233-A1.
XX
PD 24-JUN-1999.
XX
XX PF 15-DEC-1998; 98WO-GB003766.
XX PR 16-DEC-1997; 97GB-00026539.
XX
XX (UYDU-) UNIV DUNDEE.
XX PA Schor SL, Schor AM;
XX PI
XX WPI; 1999-430039/36.
XX
PT Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.
XX
XX Disclosure; Fig 2; 86pp; English.
XX
CC The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. The present sequence
CC represents the human fibronectin
XX
SQ Sequence 720 AA;

Query Match 100.0%; Score 627; DB 2; Length 720;
Best Local Similarity 100.0%; Pred. No. 7e-55; Indels 0; Gaps 0;
Matches 110; Conservative 0; Mismatches 0;
QY 1 PIAKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
|||||
Db 199 PIAKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 258
|||||
QY 61 RIGDTSKDKNRGNLLQICITGNGRGEWKRCERTSVQTTSSGSGPFTDVR 110
|||||
Db 259 RIGDTSKDKNRGNLLQICITGNGRGEWKRCERTSVQTTSSGSGPFTDVR 308

RESULT 5

ADP75952
ID ADP75952 standard; protein; 1179 AA.
AC ADP75952;
XX
XX 09-SEP-2004 (first entry)
DT
XX Human mini fibronectin protein SeqID2.
DE
XX matrix binding region; cell surface; extracellular matrix;
KW leukaemia inhibitory factor region; cell differentiation;

KW embryonic stem cell; human; mini fibronectin.

OS Homo sapiens.
XX
XX JP2004166641-A.
XX
XX 17-JUN-2004.
XX
XX 21-NOV-2002; 2002JP-00338373.
XX
XX 21-NOV-2002; 2002JP-00338373.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX (AIZU/) AIZU Y.

DR WPI; 2004-445579/42.
XX N-PSDB; ADP75954.

XX Novel protein which has matrix binding region which assembles to cell
PT surface, or to extracellular matrix surrounding cell and leukemia
PT inhibitory factor region that suppresses cell differentiation of
PT embryonic stem cell.
XX

PS Disclosure; SEQ ID NO 2; 49pp; Japanese.

XX This invention relates to a novel protein (and the gene which encodes it)
CC which has a matrix binding region which assembles to a cell surface, or
CC extracellular matrix surrounding the cell, and a leukaemia inhibitory
CC factor region which suppresses cell differentiation of embryonic stem
CC cells. The invention is useful for suppressing cell differentiation. The
CC invention can be used for the efficient inhibition of embryonic stem cell
CC differentiation. The invention thus enables efficient and continuous
CC proliferation of embryonic stem cells. Therefore, embryonic stem cells
CC can be prepared in large quantities and utilised for fundamental research
CC and applications. The present sequence is that of the human mini-
CC fibronectin protein which was used in the exemplification of the
XX invention.

SQ Sequence 1179 AA;

Query Match 100.0%; Score 627; DB 8; Length 1179;
Best Local Similarity 100.0%; Pred. No. 1.2e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PIAKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
|||||
Db 150 PIAKCFDHAAGTSYVVGTEWKEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 209
|||||

QY 61 RIGDTSKDKNRGNLLQICITGNGRGEWKRCERTSVQTTSSGSGPFTDVR 110
|||||
Db 210 RIGDTSKDKNRGNLLQICITGNGRGEWKRCERTSVQTTSSGSGPFTDVR 259

RESULT 6

ADP75957
ID ADP75957 standard; protein; 1359 AA.
XX
XX ADP75957;
AC
XX 09-SEP-2004 (first entry)
DT
XX Human leukaemia inhibiting factor/mini fibronectin protein SeqID7.
DE
XX matrix binding region; cell surface; extracellular matrix;
KW leukaemia inhibitory factor region; cell differentiation;
KW embryonic stem cell; human.
XX
XX Homo sapiens.
XX OS
XX Chimeric.
XX
XX JP2004166641-A.
XX
XX 17-JUN-2004.
PD

XX 21-NOV-2002; 2002JP-00338373.
PF XX
XX 21-NOV-2002; 2002JP-00338373.
PR XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
PA (AIZU/) AIZU Y.
PA XX
XX WPI; 2004-445579/42.
XX N-PSDB; ADP75958.
DR
XX
XX Novel protein which has matrix binding region which assembles to cell
PT surface, or to extracellular matrix surrounding cell and leukemia
PT inhibitory factor region that suppresses cell differentiation of
PT embryonic stem cell.
XX
XX Claim 11; SEQ ID NO 7; 49pp; Japanese.
PS
XX This invention relates to a novel protein (and the gene which encodes it)
XX which has a matrix binding region which assembles to a cell surface, or
CC extracellular matrix surrounding the cell, and a leukemia inhibitory
CC factor region which suppresses cell differentiation of embryonic stem
CC cells. The invention is useful for suppressing cell differentiation. The
CC invention can be used for the efficient inhibition of embryonic stem cell
CC differentiation. The invention thus enables efficient and continuous
CC proliferation of embryonic stem cells. Therefore, embryonic stem cells
CC can be prepared in large quantities and utilised for fundamental research
CC and applications. The present sequence is that of the protein of the
CC invention.
XX
SQ Sequence 1359 AA;
Query Match 100.0%; Score 627; DB 8; Length 1359;
Best Local Similarity 100.0%; Pred. NO. 1.4e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAEKFDAAGTSYVVGTEKPYQGMWVDCTCLGEGSGRITCTSRNRNDQDRTSY 60
Db 150 PIAEKFDAAGTSYVVGTEKPYQGMWVDCTCLGEGSGRITCTSRNRNDQDRTSY 209
Qy 61 RIGDTWSKKNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 210 RIGDTWSKKNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPPTDVR 259
RESULT 7
AB001289
ID ABO01289 standard; protein; 2220 AA.
AC ABO01289;
XX
XX 06-AUG-2003 (first entry)
XX Human protein NOV1b.
XX
KW Human; NOVX; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; prostate cancer; diabetes; metabolic disorder;
KW neoplasm; graft versus host disease; AIDS; bronchial asthma;
KW Crohn's disease; multiple sclerosis; infectious disease; anorexia;
KW cancer-associated cachexia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; wasting disorder; gene therapy;
KW SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 639 /note= "May be Phe as the result of a single nucleotide
FT polymorphism"
XX
XX WO2003023008-A2.
XX

PD 20-MAR-2003.
XX
XX 09-SEP-2002; 2002WO-US028596.
XX
XX 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 12-SEP-2001; 2001US-0318765P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324990P.
PR 15-FEB-2002; 2002US-0357303P.
PR 28-FEB-2002; 2002US-0360973P.
PR 20-MAR-2002; 2002US-0366131P.
PR 25-MAR-2002; 2002US-0367753P.
PR 02-APR-2002; 2002US-0369479P.
PR 10-MAY-2002; 2002US-0379532P.
PR 17-MAY-2002; 2002US-0381664P.
PR 17-MAY-2002; 2002US-0381672P.
PR 28-MAY-2002; 2002US-0383651P.
PR 29-MAY-2002; 2002US-0384012P.
PR 19-JUN-2002; 2002US-0390155P.
PR 06-SEP-2002; 2002US-00390155.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Zhong M, Li L, Gorman L, Spytek KA, Kekuda R, Taupier RJ;
PI Anderson DW, Vernet CAM, Catterton E, Miller CE, Shenoy SG;
PI Patturajan M, Pena CE, Tchernev VT, Padigar M, Gusev VF;
PI Malyankar UM, Burgess CB, Gerlach VL, Casman SJ, Rieger DK;
PI Grose WW, Smithson G, Peyman JA, Starling G, Rothenberg ME;
PI Larochelle WJ, Shinkets RA, Crabtree J, Rastelli L, Voss EZ;
PI Boldog FL, Edinger SR, Millet I, Macdougall JR, Ellerman K;
PI Chapoval A;
XX
XX WPI; 2003-313246/30.
DR N-PSDB; ACD06170.
XX
XX New polypeptides and polynucleotides having properties related to
PT stimulation of biochemical or physiological responses in a cell or
PT tissue, useful for diagnosing or preventing e.g. atherosclerosis,
PT hypertension, prostate cancer.
XX
XX Claim 2; Page 110-111; 849pp; English.
XX
XX The invention relates to an isolated polypeptide comprising one of 127
CC sequences (appearing as ABO1288-ABO1414) designated as NOVX, a mature
CC form of NOVX, an amino acid sequence which is at least 95% identical to
CC NOVX or an amino acid sequence comprising one or more conservative
CC substitutions in NOVX. Also included are nucleic acids encoding NOVX
CC proteins, determining the presence or amount of NOVX or NOVX DNA in a
CC sample (by introducing the sample to an antibody that binds
CC immunospecifically to the polypeptide, and determining the presence or
CC amount of antibody bound to the polypeptide), determining the presence of
CC or predisposition to a disease associated with altered levels of
CC expression of NOVX or NOVX DNA in a first mammalian subject, identifying
CC an agent that binds to NOVX, identifying a potential therapeutic agent
CC for treatment of a pathology related to aberrant expression or aberrant
CC physiological interactions of NOVX, screening for a modulator of activity
CC of or of latency or predisposition to a pathology associated with NOVX, a
CC vector comprising NOVX DNA, a cell comprising the vector (used to produce
CC NOVX) and an anti-NOVX antibody. The NOVX nucleic acids and polypeptides
CC are useful as a marker for cell or tissue type, and in diagnosing and
CC treating pathologies, diseases, conditions or disorders associated with
CC NOVX sequences, including cardiomyopathy, atherosclerosis, hypertension,
CC congenital heart defects, prostate cancer, diabetes, metabolic disorders,
CC neoplasm, graft versus host disease, AIDS, bronchial asthma, Crohn's
CC disease, multiple sclerosis, infectious diseases, anorexia, cancer-

CC associated cachexia, neurodegenerative disorders (e.g. Alzheimer's
CC disease or Parkinson's disease), immune disorders, haematopoietic
CC disorders, dyslipidaemias, and wasting disorders associated with chronic
CC diseases. These may also be used to screen for molecules which inhibit or
CC enhance NOXV activity or function, and for detecting specific cell types.
CC These may also be used in chromosome mapping, gene therapy, tissue
CC typing, and in forensic biology. The present sequence represents a NOXV
CC protein

XX Sequence 2220 AA;

Query Match 100.0%; Score 627; DB 6; Length 2220;
Best Local Similarity 100.0%; Pred. No. 2.4e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMWMDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGTEWKPQGMWMDCTCLGEGSGRITCTSRNRCNDQDRTSY 240
Qy 61 RIGDTWSKKNRGNLLQICTGNGRGWKCERHTSVQTTSSGSGPFTDVR 110
Db 241 RIGDTWSKKNRGNLLQICTGNGRGWKCERHTSVQTTSSGSGPFTDVR 290

RESULT 8

ID ADN95950 standard; protein; 2220 AA.

XX ADN95950;

DT 01-JUL-2004 (first entry)

XX Human NOXV polypeptide #2.

KW Human; NOXV; metabolic disorder; diabetes; obesity; infectious disease;
KW anorexia; cancer; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW antidiabetic; anorectic; antimicrobial; anabolic; eating disorder;
KW cytostatic; neuroprotective; nootropic; antiparkinsonian; antianemic.

XX Homo sapiens.

OS US2004067490-A1.

XX 08-APR-2004.

XX 06-SEP-2002; 2002US-00236392.

XX 07-SEP-2001; 2001US-0318120P.

XX 07-SEP-2001; 2001US-0318130P.

XX 07-SEP-2001; 2001US-0318219P.

XX 10-SEP-2001; 2001US-0318430P.

XX 12-SEP-2001; 2001US-0318765P.

XX 17-SEP-2001; 2001US-0322781P.

XX 17-SEP-2001; 2001US-0322816P.

XX 19-SEP-2001; 2001US-0323519P.

XX 20-SEP-2001; 2001US-0323631P.

XX 20-SEP-2001; 2001US-0323636P.

XX 25-SEP-2001; 2001US-0324969P.

XX 25-SEP-2001; 2001US-0325091P.

XX 26-SEP-2001; 2001US-0324900P.

XX 15-FEB-2002; 2002US-0357303P.

XX 28-FEB-2002; 2002US-0360973P.

XX 20-MAR-2002; 2002US-0366131P.

XX 25-MAR-2002; 2002US-0367753P.

XX 02-APR-2002; 2002US-0369479P.

XX 10-MAY-2002; 2002US-0379532P.

XX 17-MAY-2002; 2002US-0381664P.

XX 17-MAY-2002; 2002US-0381672P.

XX 28-MAY-2002; 2002US-0383651P.

XX 29-MAY-2002; 2002US-0384012P.

XX 19-JUN-2002; 2002US-0390155P.

PA (ZHON/) ZHONG M.
PA (LILL/) LI L.
PA (GORM/) GORMAN L.
PA (SPYT/) SPYTEK K A.
PA (KEKU/) KEKUDA R.
PA (TAUF/) TAUFIER R J.
PA (ANDE/) ANDERSON D W.
PA (VERN/) VERNET C A M.
PA (CATT/) CATTERTON E.
PA (MILL/) MILLER C E.
PA (SHEN/) SHENOY S G.
PA (PATT/) PATTURAJAN M.
PA (PENA/) PENNA C E A.
PA (TCHE/) TCHERNEV V T.
PA (PADI/) PADIGARU M.
PA (GUSE/) GUSEV V Y.
PA (MALY/) MALYANKAR U M.
PA (BURG/) BURGESS C E.
PA (GERL/) GERLACH V.
PA (CASM/) CASMAN S J.
PA (RIEG/) RIEGER D K.
PA (GROS/) GROSSE W M.
PA (SMIT/) SMITHSON G.
PA (PEYM/) PEYMAN J A.
PA (STAR/) STARLING G.
PA (ROTH/) ROTHENBERG M E.
PA (LARO/) LAROCHELLE W J.
PA (SHIM/) SHIMKETS R A.
PA (CRAB/) CRABTREE J.
PA (RAST/) RASTELLI L.
PA (VOSS/) VOSS E Z.
PA (BOLD/) BOLDOG F L.
PA (EDIN/) EDINGER S R.
PA (MILL/) MILLET I.
PA (MACD/) MACDOUGALL J R.
PA (ELLE/) ELLERMAN K.
PA (CHAP/) CHAPOVAL A.

XX Zhong M, Li L, Gorman L, Spytek KA, Kekuda R, Taupier RJ;
PI Anderson DW, Vernet CAM, Catterton E, Miller CE, Shenoy SG;
PI Patturajan M, Pena CEA, Tchernev VT, Padigar M, Gusev VY;
PI Malynkar UM, Burgess CE, Gerlach V, Casman SJ, Rieger DK;
PI Grosse WM, Smithson G, Peyman JA, Starling G, Rothenberg ME;
PI Larochelle WJ, Shimkets RA, Crabtree J, Rastelli L, Voss EZ;
PI Boldog FL, Edinger SR, Millet I, Macdougall JR, Ellerman K;
PI Chapoval A;

XX MPI; 2004-355290/33.

XX N-PSDB; ADN95949.

XX New isolated polypeptide, useful for treating or preventing a pathology
XX associated with the polypeptide, e.g. diabetes, infectious disease,
XX cancer, neurodegenerative disorders or Alzheimer's disease.

XX Claim 2; SEQ ID NO 4; 552pp; English.

XX The invention relates to human NOXV polypeptides and polynucleotides. The
XX isolated nucleic acids can be used to express the novel proteins, to
XX detect novel mRNA or a genetic lesion in a novel gene and to modulate its
XX activity. It can also be used in gene therapy for treating or preventing
XX a pathology associated with the protein or nucleic acid. The disorders
XX include metabolic disorders, diabetes, obesity, infectious diseases,
XX anorexia, cancer, neurodegenerative disorders, Alzheimer's disease,
XX Parkinson's disease, immune disorders and haematopoietic disorders. This
XX sequence represents a human NOXV polypeptide of the invention.

XX Sequence 2220 AA;

Query Match 100.0%; Score 627; DB 8; Length 2220;
Best Local Similarity 100.0%; Pred. No. 2.4e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMWMDCTCLGEGSGRITCTSRNRCNDQDRTSY 60

Db 181 PIAKCFDHAAGTSYVVGGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDRTSY 240
QY 61 RIGDWTWSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
Db 241 RIGDWTWSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290

RESULT 9
AAM38647
ID AAM38647 standard; protein; 2265 AA.
XX AC AAM38647;
XX
DT 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 1792.
XX
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX
XX 21-JAN-2000; 2000US-00488725.
XX
XX 25-APR-2000; 2000US-00552317.
XX
XX 20-JUN-2000; 2000US-00598042.
XX
XX 19-JUL-2000; 2000US-00620312.
XX
XX 03-AUG-2000; 2000US-00653450.
XX
XX 14-SEP-2000; 2000US-00662191.
XX
XX 19-OCT-2000; 2000US-00693036.
XX
XX 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang JT, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Dermanac RT;
XX
XX WPI; 2001-442253/47.
XX
XX N-PSDB; AAI57803.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 3; SEQ ID NO 1792; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

SQ Sequence 2265 AA;
Query Match 100.0%; Score 627; DB 4; Length 2265;
Best Local Similarity 100.0%; Pred. No. 2.5e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PIAKCFDHAAGTSYVVGGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDRTSY 240
QY 61 RIGDWTWSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
Db 241 RIGDWTWSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290

RESULT 10
ABR40124
ID ABR40124 standard; protein; 2266 AA.
XX AC ABR40124;
XX
DT 04-JUL-2003 (first entry)
XX
XX Human cell adhesion and extracellular matrix protein, CADECM-21.
XX
XX Human; anti-HIV; antiallergic; cerebroprotective; antiparkinsonian;
KW anticonvulsant; nontropic; neuroprotective; immunosuppressive;
KW dermatological; anti-inflammatory; cytostatic; antiarteriosclerotic;
KW gene therapy; cell adhesion; extracellular matrix; CADECM;
KW immune system disorder; AIDS; allergy; neurological disorder; stroke;
KW Parkinson's disease; epilepsy; developmental disorder; Down's syndrome;
KW cerebral palsy; connective tissue disorder; systemic lupus erythematosus;
KW genetic disorder; Alport's syndrome; cell proliferative disorder; cancer;
KW atherosclerosis.
XX
XX Homo sapiens.
XX
XX WO2003027230-A2.
XX
XX 03-APR-2003.
XX
XX 02-AUG-2002; 2002WO-US024649.
XX
XX 03-AUG-2001; 2001US-0309964P.
XX
XX 03-AUG-2001; 2001US-0310119P.
XX
XX 17-AUG-2001; 2001US-0313091P.
XX
XX 31-AUG-2001; 2001US-0316771P.
XX
XX 07-SEP-2001; 2001US-0317896P.
XX
XX 21-SEP-2001; 2001US-0324781P.
XX
XX 05-OCT-2001; 2001US-0327606P.
XX
XX 12-OCT-2001; 2001US-0328960P.
XX
XX 09-NOV-2001; 2001US-034471P.
XX
XX 17-MAY-2002; 2002US-0381291P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Burford N, Warren BA, Duggan BM, Mason PM, Richardson TW, Yue H;
PI Forsythe IJ, Elliott VS, Griffin JA, Gorvad AE, Azimzai Y;
PI Kallick DA, Xu Y, Honchell CD, Baughn MR, Gietzen KJ, Lee S;
PI Wallia NK, Tang YT, Nguyen DB, Becha SD, Lee SY, Ramkumar J;
XX
XX WPI; 2003-354645/33.
XX
XX N-PSDB; ACC00412.
XX
XX New human cell adhesion and extracellular matrix proteins (CADECM),
PT useful for diagnosing, treating or preventing disorders associated with
PT aberrant CADECM expression, e.g. cancer, AIDS, atherosclerosis, allergies
PT or stroke.
XX
XX Claim 1; Page 203-209; 234pp; English.
XX
XX The present invention relates to novel human cell adhesion and
CC extracellular matrix proteins, CADECM-1 to CADECM-22, and their coding

The invention relates to selective tissue vascular thrombogen (STVT) protein comprising a selective binding domain associated with a tissue factor (TF) polypeptide, where the selective binding domain can bind to a channel for blood within a tissue and the human tissue factor can initiate thrombosis within the channel. STVT is useful for treating a solid tumour in an animal. Pharmaceutical preparation comprising STVT is useful for the manufacture of a medicament for treating solid tumour such as lung, breast, ovary, stomach, pancreas, larynx, oesophagus, testes, liver, prostatic, biliary tract, colon, rectum, cervix, uterus.

CC bind to channels within any tissue,
CC binding, the tissue factor polypeptide
CC used in a claimed method of treating

CC a lung, breast, ovary, stomach, pancreas, larynx, oesophagus, testis,
CC liver, parotid, biliary tract, colon, rectum, cervix, uterus.
CC endometrium, kidney, bladder, prostate, thyroid, benign prostate
CC hyperplasia, squamous cell carcinoma, adenocarcinoma, small cell
CC carcinoma, melanoma, glioma or neuroblastoma tumour, and where thrombosis
CC leads to tumour necrosis
XX
SQ Sequence 2320 AA;

Query Match 100.0%; Score 627; DB 6; Length 2320;
Best Local Similarity 100.0%; Pred. No. 2.6e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGSTWKPQYQWMMVDTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 146 PIAKCFDHAAGTSYVVGSTWKPQYQWMMVDTCLGEGSGRITCTSRNRCNDQDRTSY 205
QY 61 RIGDTWSKKDNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 206 RIGDTWSKKDNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPPTDVR 255

RESULT 13
ABO01288
ID ABO01288 standard; protein; 2320 AA.

AC ABO01288;

XX 07-AUG-2003 (first entry)

XX Human protein NOV1a.

XX Human; NOVX; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; prostate cancer; diabetes; metabolic disorder;
KW neoplasm; graft versus host disease; AIDS; bronchial asthma;
KW Crohn's disease; multiple sclerosis; infectious disease; anorexia;
KW cancer-associated cachexia; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; wasting disorder; gene therapy.

XX Homo sapiens.

XX WO2003023008-A2.

XX 20-MAR-2003.

XX 09-SEP-2002; 2002WO-US028596.

XX 07-SEP-2001; 2001US-0318120P.

XX 07-SEP-2001; 2001US-0318130P.

XX 10-SEP-2001; 2001US-0318430P.

XX 12-SEP-2001; 2001US-0318765P.

XX 17-SEP-2001; 2001US-0322781P.

XX 17-SEP-2001; 2001US-0322816P.

XX 19-SEP-2001; 2001US-0323519P.

XX 20-SEP-2001; 2001US-0323631P.

XX 20-SEP-2001; 2001US-0323636P.

XX 25-SEP-2001; 2001US-0324969P.

XX 25-SEP-2001; 2001US-0325091P.

XX 26-SEP-2001; 2001US-0324990P.

XX 15-FEB-2002; 2002US-0357303P.

XX 28-FEB-2002; 2002US-0360973P.

XX 20-MAR-2002; 2002US-0366131P.

XX 25-MAR-2002; 2002US-0367753P.

XX 02-APR-2002; 2002US-0369479P.

XX 10-MAY-2002; 2002US-0379532P.

XX 17-MAY-2002; 2002US-0381664P.

XX 17-MAY-2002; 2002US-0381672P.

XX 28-MAY-2002; 2002US-0383651P.

XX 29-MAY-2002; 2002US-0384012P.

XX 19-JUN-2002; 2002US-0390155P.

XX 06-SEP-2002; 2002US-00390155.

(CURA-) CURAGEN CORP.

XX Zhong M, Li L, Gorman L, Spytek KA, Kekuda R, Taupier RJ;
PI Anderson DW, Vernet CAM, Catterton E, Miller CE, Shenoy SG;
PI Patturajan M, Pena CEA, Tchernev VI, Padigan M, Gusev VI;
PI Malyankar UN, Burgess CE, Gerlach VL, Casman SJ, Rieger DK;
PI Grosse WM, Smithson G, Peyman JA, Starling J, Rothenberg ME;
PI Larochele WJ, Shinkets RA, Crabtree J, Rastelli L, Voss EZ;
PI Boldog FL, Edinger SR, Millet I, Macdougall JR, Ellerman K;
PI Chapoval A;

XX WPI; 2003-313246/30.

DR N-PSDB; ACD06169.

XX New polypeptides and polynucleotides having properties related to
PT stimulation of biochemical or physiological responses in a cell or
PT tissue, useful for diagnosing or preventing e.g. atherosclerosis,
PT hypertension, prostate cancer.

XX Claim 2; Page 107-108; 849pp; English.

XX The invention relates to an isolated polypeptide comprising one of 127
CC sequences (appearing as ABO1288-ABO1414), designated as NOVX, a mature
CC form of NOVX, an amino acid sequence which is at least 95% identical to
CC NOVX or an amino acid sequence comprising one or more conservative
CC substitutions in NOVX. Also included are nucleic acids encoding NOVX
CC proteins, determining the presence or amount of NOVX or NOVX DNA in a
CC sample (by introducing the sample to an antibody that binds
CC immunospecifically to the polypeptide, and determining the presence or
CC amount of antibody bound to the polypeptide), determining the presence of
CC or predisposition to a disease associated with altered levels of
CC expression of NOVX or NOVX DNA in a first mammalian subject, identifying
CC an agent that binds to NOVX, identifying a potential therapeutic agent
CC for treatment of a pathology related to aberrant expression or aberrant
CC physiological interactions of NOVX, screening for a modulator of activity
CC of or of latency or predisposition to a pathology associated with NOVX, a
CC vector comprising NOVX DNA, a cell comprising the vector (used to produce
CC NOVX) and an anti-NOVX antibody. The NOVX nucleic acids and polypeptides
CC are useful as a marker for cell or tissue type, and in diagnosing and
CC treating pathologies, diseases, conditions or disorders associated with
CC NOVX sequences, including cardiomyopathy, atherosclerosis, hypertension,
CC congenital heart defects, prostate cancer, diabetes, metabolic disorders,
CC neoplasm, graft versus host disease, AIDS, bronchial asthma, Crohn's
CC disease, multiple sclerosis, infectious diseases, anorexia, cancer-
CC associated cachexia, neurodegenerative disorders (e.g. Alzheimer's
CC disease or Parkinson's disease), immune disorders, haematopoietic
CC disorders, dyslipidaemias, and wasting disorders associated with chronic
CC diseases. These may also be used to screen for molecules which inhibit or
CC enhance NOVX activity or function, and for detecting specific cell types.
CC These may also be used in chromosome mapping, gene therapy, tissue
CC typing, and in forensic biology. The present sequence represents a NOVX
CC protein

XX SQ Sequence 2320 AA;

Query Match 100.0%; Score 627; DB 6; Length 2320;
Best Local Similarity 100.0%; Pred. No. 2.6e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGSTWKPQYQWMMVDTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 146 PIAKCFDHAAGTSYVVGSTWKPQYQWMMVDTCLGEGSGRITCTSRNRCNDQDRTSY 205
QY 61 RIGDTWSKKDNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 206 RIGDTWSKKDNRGNLLQCICTGNRGWKCERHTSVQTTSSGSGPPTDVR 255

RESULT 14

AAR92778

ID AAR92778 standard; protein; 2324 AA.

XX AAR92778;

AC AAR92778;

XX DT 21-JUN-1996 (first entry)
XX DE Human fibronectin.
XX KW Fibrin-binding peptide; fibronectin; fibrinolysis; fibrinogenesis;
XX KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis; therapy.
XX OS Homo sapiens.
XX FH Key
XX FT Location/Qualifiers
XX FT 21..65
XX FT /label= 1F1
XX FT /note= "type 1 module 1"
XX FT 66..109
XX FT /label= 2F1
XX FT /note= "type 1 module 2"
XX FT 110..154
XX FT /label= 3F1
XX FT /note= "type 1 module 3"
XX FT 155..199
XX FT /label= 4F1
XX FT /note= "type 1 module 4"
XX FT 200..244
XX FT /label= 5F1
XX FT /note= "type 1 module 5"
XX FT 277..312
XX FT /label= 6F1
XX FT /note= "type 1 module 6"
XX FT 439..486
XX FT /label= 7F1
XX FT /note= "type 1 module 7"
XX FT 439..467
XX FT /note= "Cys439-Cys467 disulfide bond"
XX FT 465..477
XX FT /note= "Cys465-Cys477 disulfide bond"
XX FT 487..529
XX FT /label= 8F1
XX FT /note= "type 1 module 8"
XX FT 530..569
XX FT /label= 9F1
XX FT /note= "type 1 module 9"
XX FT 2123..2230
XX FT /label= Fibrin binding domain
XX FT /note= "11 kDa C-terminal fibrin-binding domain"
XX FT 2141..2185
XX FT /label= 10F1
XX FT /note= "type 1 module 10"
XX FT 2144..2173
XX FT /note= "Cys2144-Cys2173 disulfide bond"
XX FT 2171..2183
XX FT /note= "Cys2171-Cys2183 disulfide bond"
XX FT 2187..2230
XX FT /label= 11F1
XX FT /note= "type 1 module 11"
XX FT 2189..2216
XX FT /note= "Cys2189-Cys2216 disulfide bond"
XX FT 2214..2226
XX FT /note= "Cys2214-Cys2226 disulfide bond"
XX FT 2233..2271
XX FT /label= 12F1
XX FT /note= "type 1 module 12"
XX PN W09604304-A1.
XX PD 15-FEB-1996.
XX PF 01-AUG-1995; 95WO-US009819.
XX PR 01-AUG-1994; 94US-00283857.
XX PA (UUNY) UNIV NEW YORK STATE.
XX

PI Gold LI, Rostagno AA;
XX WPI; 1996-129333/13.
XX DR N-PSDB; AAT17551.
XX PT New fibrin-binding peptide molecules - used for the diagnosis and
XX PT treatment of conditions associated with fibrin deposition, e.g. thrombi.
XX PS Claim 2; Page 104-110; 146pp; English.
XX CC Human fibronectin (AAR92778) has a fibrin-binding site, close to the C-
XX CC terminus (amino acids 2123-232 or 2141-2320), covering the 10F1.11F1
XX CC module pair of the protein. Fibrin-binding sites can be prep'd. by
XX CC enzymatic cleavage of fibronectin, peptide synthesis or by recombinant
XX CC DNA techniques. They are used to detect a fibrin-binding target site, to
XX CC treat disorders involving abnormal fibrinolysis or fibrinogenesis, to
XX CC disrupt blood clots and to aid delivery of medicaments to fibrin- contg.
XX CC sites. They show high affinity to and slow dissociation from fibrin, and
XX CC provide fast diffusion and rapid clearance
XX SQ Sequence 2324 AA;
Query Match 100.0%; Score 627; DB 2; Length 2324;
Best Local Similarity 100.0%; Pred. NO. 2.6e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PIAKCFDHAAGTSYVVGTEWKPQGMWMDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
DB 150 PIAKCFDHAAGTSYVVGTEWKPQGMWMDCTCLGEGSGRITCTSRNRCNDQDRTSY 209
QY 61 RIGDTWSKDNRGNLQICITGNRGGEWKERHTSVQTTSSGSGPFTDVR 110
DB 210 RIGDTWSKDNRGNLQICITGNRGGEWKERHTSVQTTSSGSGPFTDVR 259
RESULT 15
AAU74674
ID AAU74674 standard; protein; 2324 AA.
XX AC AAU74674;
XX DT 09-APR-2002 (first entry)
XX DE Human fibronectin protein.
XX KW Human; fibronectin; VLA-4 binding site; very late antigen-4;
XX KW quiescent haematopoietic cell; HC; apoptosis; CD34+.
XX OS Homo sapiens.
XX PN W0200187071-A1.
XX PD 22-NOV-2001.
XX PF 12-MAY-2000; 2000WO-US012993.
XX PR 12-MAY-2000; 2000WO-US012993.
XX PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.
XX PI Williams DA, Bradford GB, Dutt P, Yoder MC;
XX WPI; 2002-082932/11.
XX DR Obtaining hematopoietic cell population containing quiescent cells for
XX PT use in treating a subject, by expanding the cells while adhered to a
XX PT polypeptide containing binding sites for integrins on the cells.
XX PS Disclosure; Page 51-61; 63pp; English.
XX CC The invention relates to obtaining a population of quiescent
XX CC hematopoietic cells (HC), comprising culturing HC while adhering the
XX CC cells to a polypeptide having a very late antigen-4 (VLA-4) binding site,

in particular fibronectin polypeptide, so as to expand the number of HC, where the adhering provides an increased percentage of quiescent HC. Also included are inducing apoptosis of a subpopulation of HC, by contacting the cells with a polypeptide having a VLA-4 binding site under conditions to cause apoptosis of a subpopulation of HC and a medium for culturing HC which enriches quiescent HC, comprising a fibronectin polypeptide. The method is useful for obtaining a cell population containing quiescent haematopoietic cells which are useful for treating a subject. Use of the VLA-4 containing peptide is useful for inducing apoptosis of a subpopulation of CD34+ HC population. The present sequence is human fibronectin which contains then VLA-4 binding site used in the method of the invention

XX
SQ Sequence 2324 AA;

```
Query Match      100.0%; Score 627; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2.6e-54;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 PIAEKCFDHAAGTSYVVGETWEKPYQGWMVVDCTCLGEGSGRI TCTSRNRCNDQDTRTSY 60

db 150 PIAEKCFDHAAGTSYVVGETWEKPYOGWMMVDCTCLGEGSGRITCTSRNRNCNDQDTRTSY 209

61 RIGDTWSKKDNRGNLLOCICTGNRGEWK CERHTSVOTSSGSGPFTDVR 110 QV

Db 210 RIGDTWSKDNRGNLLOCICTGNRGGEWK CERHTSVOTTSSSGPFTDVR 259

Search completed: November 3, 2004, 23:49:26
Job time : 54.4931 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:42:10 ; Search time 11.3014 Seconds
(without alignments)
936.510 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAKCFDPAAGTSYVGET.....ERTSVQTTSSGSGPFDVR 110

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	627	100.0	2386	1 FNHU	fibronectin precu
2	605	96.5	2265	1 FNBO	fibronectin - bovi
3	580.5	92.6	2477	2 S14428	fibronectin precu
4	539	86.0	2481	2 A43908	fibronectin - Afri
5	192	30.6	1020	2 A29355	fibronectin - chic
6	173.5	27.7	190	2 I51279	fibronectin - east
7	97.5	15.6	2476	2 T34022	zonadhesin - pig
8	83.5	13.3	103	2 A49173	fibronectin - mous
9	75	12.0	1090	2 A41696	regulatory protein
10	74.5	11.9	3097	2 T00021	DN-cadherin - fru
11	71.5	11.4	495	2 B71360	hypothetical prote
12	71	11.3	810	2 T10756	Nel-homolog protei
13	70.5	11.2	773	2 I45059	beta-1 integrin su
14	70.5	11.2	1074	2 T04777	hypothetical prote
15	70.5	11.2	1169	2 T18423	hypothetical prote
16	70	11.2	719	2 T00266	hypothetical prote
17	70	11.2	788	2 T25061	hypothetical prote
18	68.5	10.9	550	2 T47158	hypothetical prote
19	68.5	10.9	5376	2 T42215	zonadhesin - mous
20	68	10.8	385	2 T26487	hypothetical prote
21	68	10.8	646	2 JN0473	P-selectin precurs
22	68	10.8	863	1 S51789	VLDR receptor prec
23	68	10.8	2352	2 T30201	Notch homolog prot
24	67.5	10.8	1372	2 T25933	hypothetical prote
25	67.5	10.8	3461	2 S58870	reelin precursor -
26	67	10.7	340	2 T32104	hypothetical prote
27	66.5	10.6	346	2 E95870	probable dtdppgluc
28	66.5	10.6	429	2 S23581	lamB protein precu
29	66.5	10.6	509	2 A45992	activin A receptor

RESULT 1

FNHU

fibronectin precursor [validated] - human
N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 09-Jul-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A2

R;Dean, D.C.; Bowls, C.L.; Bourgeois, S

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656

A;Accession: A26460

A;Molecule type: DNA

A;Residues: 1-49 <DE>

A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; GB:M15801; NID:gl82686; PIDN:AAA533

R;Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A;Title: Evolution of the fibronectin gene.

A;Reference number: A26284; MUID:86111901; PMID:3003095

A;Accession: A26284

A;Molecule type: DNA

A;Residues: 1447-1540 <OLD>

A;Cross-references: GB:M12549; NID:gl82688

A;Note: the authors translated the codon TTC for residue 1494 as Glu

R;Paolella, G.; Henchcliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-

A;Reference number: S00848; MUID:88233940; PMID:3375063

A;Accession: S03917

A;Molecule type: DNA

A;Residues: 1594-1767, 'V', 1769-1783 <PAO>

A;Cross-references: EMBL:X07718; NID:g31402

A;Note: the authors translated the codon AAC for residue 1631 as Asp

R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene

A;Reference number: A24854; MUID:87030929; PMID:3770201

A;Accession: A24854

A;Molecule type: DNA

A;Residues: 1992-2147 <VIB>

A;Cross-references: GB:X04530; NID:g31436

R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.

FEBS Lett. 207, 145-148, 1986

A;Title: Human fibronectin is synthesized as a pre-propolypeptide.

A;Reference number: A24476; MUID:87030890; PMID:3770189

A;Accession: A24476

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-14, 'Q', 16-38 <GUT>

R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1755-1759, 1985

A;Title: Primary structure of human fibronectin: differential splicing may generate at

A;Reference number: A31008; MUID:85284965; PMID:2992939
A;Accession: A91008
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-1344,1346-2080;2112-2386 <KOR>
A;Cross-references: GB:X02761
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Nucleic Acids Res. 12, 5853-5868, 1984
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptides
A;Reference number: A93529; MUID:84272258; PMID:6462919
A;Accession: A93529
A;Molecule type: mRNA
A;Residues: 773-2080;2112-2386 <K02>
A;Cross-references: GB:X00739
R;Oldberg, A.; Linney, E.; Ruoslahti, E.
J. Biol. Chem. 258, 10193-10196, 1983
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell
A;Reference number: A21011; MUID:83290929; PMID:6688418
A;Accession: A21011
A;Molecule type: mRNA
A;Residues: 1434-1537 <OL2>
A;Cross-references: GB:K00055; NID:g182680; PIDN:AAAS2459.1; PID:g182683
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.
Biochemistry 24, 2698-2704, 1985
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with rat
A;Reference number: A90495; MUID:85280409; PMID:2992573
A;Accession: A90495
A;Molecule type: mRNA
A;Residues: 1594-2386 <BER>
A;Cross-references: GB:M10905; NID:g182696; PIDN:AAAS2462.1; PID:g182697
R;Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.
FEBS Lett. 186, 31-34, 1985
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.
A;Reference number: A22245; MUID:85231203; PMID:2989004
A;Accession: A22245
A;Molecule type: mRNA
A;Residues: 1948-2067 <UME>
A;Cross-references: GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID:g182706
A;Accession: B22245
A;Molecule type: mRNA
A;Residues: 1975-1991;2017-2039 <UM2>
R;Sekiguchi, K.; Kios, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.
Biochemistry 25, 4936-4941, 1986
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes
A;Reference number: 152394; MUID:87026578; PMID:3021206
A;Accession: I65273
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1978-1990,2016-2018,'N',2020-2081,2113-2127 <SEK>
A;Cross-references: GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID:g182704
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectins
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: A21165
A;Molecule type: mRNA
A;Residues: 2291-2386 <K03>
A;Cross-references: GB:K00799; NID:g182681; PIDN:AAAS2460.1; PID:g182684
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 258, 12670-12674, 1983
A;Title: Primary structure of human plasma fibronectin.
A;Reference number: A92398; MUID:84032463; PMID:6630202
A;Accession: A92398
A;Molecule type: protein
A;Residues: 32-47,'C',49-51,'S',53-72,'A',74-290 <GAR1>
R;Garcia-Pardo, A.; Gold, L.I.
Arch. Biochem. Biophys. 304, 181-188, 1993
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the R
A;Reference number: S34791; MUID:93312001; PMID:8323285
A;Accession: S34791
A;Molecule type: protein
A;Residues: 291-300;551-560 <GAR2>

R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.
Thromb. Res. 43, 469-477, 1986
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.
A;Reference number: A60904; MUID:87019725; PMID:3532418
A;Accession: A60904
A;Molecule type: protein
A;Residues: 293-301 <GAR>
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Sirl, A.; Shively, J.E.; Zardi, L.
J. Biol. Chem. 260, 12136-12141, 1985
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pl
A;Reference number: A23901; MUID:86008277; PMID:3900070
A;Accession: A23901
A;Molecule type: protein
A;Residues: 616-677,'Q',679-703,'PT' <CAL>
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.
J. Biol. Chem. 257, 9593-9597, 1982
A;Title: The cell attachment domain of fibronectin. Determination of the primary structu
A;Reference number: A92386; MUID:82265604; PMID:7050098
A;Accession: A92386
A;Molecule type: protein
A;Residues: 1441-1548 <PIE>
A;Note: Residues 1524-1527 are responsible for the cell-binding activity
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.
Biochem. J. 241, 923-928, 1987
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom
A;Reference number: A32517; MUID:87421275; PMID:3593230
A;Accession: A32517
A;Molecule type: protein
A;Residues: 1589-1630,'T',1722-2058 <GAR3>
R;Tresselt, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pand
Biochem. J. 274, 731-738, 1991
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A
A;Reference number: S14357; MUID:91190085; PMID:2012601
A;Accession: S14357
A;Molecule type: protein
A;Residues: 1614-1630,'T',1722-2081,2113-2244 <TRE>
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.
J. Biol. Chem. 260, 10320-10325, 1985
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dal
A;Reference number: A23891; MUID:85261459; PMID:4019516
A;Accession: A23891
A;Molecule type: protein
A;Residues: 2071-2080;2112-2356 <GAR4>
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tis
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins;
ation, and transformation.
C;Genetics:
A;Gene: GDB:FNI
A;Cross-references: GDB:119135; OMIM:135600
A;Map position: 2q34-2q34
A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II repe
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplica
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-31/Domain: propeptide #status predicted <PRO>
F;32-2386/Product: fibronectin #status experimental <MAT>
F;52-272/Domain: fibrin and heparin binding <FHB>
F;57-135/Domain: fibronectin type I repeat homology <1F1>
F;57-135/Domain: fibronectin type I repeat homology <1F2>
F;141-179/Domain: fibronectin type I repeat homology <1F3>
F;186-225/Domain: fibronectin type I repeat homology <1F4>
F;231-270/Domain: fibronectin type I repeat homology <1F5>
F;308-608/Domain: collagen binding <CBR>
F;308-342/Domain: fibronectin type I repeat homology <1F6>
F;360-401/Domain: fibronectin type II repeat homology <2F1>
F;420-461/Domain: fibronectin type II repeat homology <2F2>
F;470-508/Domain: fibronectin type I repeat homology <1F7>
F;518-555/Domain: fibronectin type I repeat homology <1F8>
F;561-599/Domain: fibronectin type I repeat homology <1F9>
F;609-692/Domain: fibronectin type III repeat homology <3FA>
F;616-706/Domain: heparin binding <HBP>
F;719-801/Domain: fibronectin type III repeat homology <3FB>
F;810-891/Domain: fibronectin type III repeat homology <3FC>

F;906-988/Domain: fibronectin type III repeat homology <3FD>
F;996-1077/Domain: fibronectin type III repeat homology <3FE>
F;1086-1164/Domain: fibronectin type III repeat homology <3FF>
F;1173-1258/Domain: fibronectin type III repeat homology <3FG>
F;1266-1349/Domain: fibronectin type III repeat homology <3FH>

Query Match 100.0%; Score 627; DB 1; Length 2386;
Best Local Similarity 100.0%; Pred. No. 1.6e-51;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHRAGTSYVVGTEWKPQGMWVDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
DB 181 PIAKCFDHRAGTSYVVGTEWKPQGMWVDCTCLGEGSGRITCTSRNRCNDQDRTSY 240
QY 61 RIGDTWSKKDNRGNLLQICITGNRGWKCERHTSVQTTSSGSGPFTDVR 110
DB 241 RIGDTWSKKDNRGNLLQICITGNRGWKCERHTSVQTTSSGSGPFTDVR 290

RESULT 2
FNBO
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Dec-1988 #sequence revision 31-Dec-1988 #text change 09-Jul-2004
C;Accession: A26452; B21165; A23292
R;Skorstengaard, K.; Jensen, M.S.; Sahl, P.; Petersen, T.E.; Magnusson, S.
Eur. J. Biochem. 161, 441-453, 1986
A;Title: Complete primary structure of bovine plasma fibronectin.
A;Reference number: A26452; MUID:87054047; PMID:3780752
A;Accession: A26452
A;Molecule type: protein
A;Residues: 1-2265 <SKO>
R;Kornbliht, A.R.; Vibe-Pedersen, K.; Baralle, F.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectin
A;Reference number: A21165; MUID:83221567; PMID:6304699
A;Accession: B21165
A;Molecule type: mRNA
A;Residues: 2170-2265 <KOR>
R;Petersen, T.E.; Thøgersen, H.C.; Skorstengaard, K.; Vibe-Pedersen, K.; Sahl, P.; Sotth
Proc. Natl. Acad. Sci. U.S.A. 80, 137-141, 1983
A;Title: Partial primary structure of bovine plasma fibronectin: three types of internal
A;Reference number: A23292; MUID:83117805; PMID:6218503
A;Accession: A23292
A;Molecule type: protein
A;Residues: 1-16,'C',18-20,'S',22-432;447-463;1367-1517;1567-1673;2062-2176,'N',2178-226
C;Comment: Cys-1201 and Cys-2015 have free sulfhydryl groups.
C;Comment: The plasma fibronectin molecule consists of two chains, which are connected b
C;Comment: Fibronectins bind cell surfaces and various compounds including collagen, fib
aling, and maintenance of cell shape.
C;Comment: Plasma fibronectin is synthesized by hepatocytes.
C;Superfamily: fibronectin; fibronectin type I repeat homology; fibronectin type II rep
C;Keywords: acute phase; alternative splicing; collagen binding; duplication; extracellu
F;21-241/Domain: fibronectin type I repeat homology <1F1>
F;21-56/Domain: fibronectin type I repeat homology <1F2>
F;66-104/Domain: fibronectin type I repeat homology <1F3>
F;110-148/Domain: fibronectin type I repeat homology <1F4>
F;155-194/Domain: fibronectin type I repeat homology <1F5>
F;200-239/Domain: fibronectin type I repeat homology <1F6>
F;277-577/Domain: collagen binding <CB>
F;277-311/Domain: fibronectin type I repeat homology <1F6>
F;329-370/Domain: fibronectin type II repeat homology <2F1>
F;389-430/Domain: fibronectin type II repeat homology <2F2>
F;439-477/Domain: fibronectin type I repeat homology <1F7>
F;487-524/Domain: fibronectin type I repeat homology <1F8>
F;530-568/Domain: fibronectin type I repeat homology <1F9>
F;578-661/Domain: fibronectin type III repeat homology <FN3A>
F;688-770/Domain: fibronectin type III repeat homology <FN3B>
F;779-860/Domain: fibronectin type III repeat homology <FN3C>
F;875-957/Domain: fibronectin type III repeat homology <FN3D>
F;965-1046/Domain: fibronectin type III repeat homology <FN3E>

F;1055-1134/Domain: fibronectin type III repeat homology <FN3F>
F;1142-1227/Domain: fibronectin type III repeat homology <FN3G>
F;1235-1318/Domain: fibronectin type III repeat homology <FN3H>
F;1326-1404/Domain: fibronectin type III repeat homology <GN3I>
F;1410-1517/Domain: cell attachment <CAD>
F;1416-1502/Domain: fibronectin type III repeat homology <FN3J>
F;1493-1495/Region: cell attachment (R-G-D) motif
F;1510-1592/Domain: fibronectin type III repeat homology <FN3K>
F;1600-1870/Domain: heparin binding <HB2>
F;1600-1682/Domain: fibronectin type III repeat homology <FN3L>
F;1692-1773/Domain: fibronectin type III repeat homology <FN3M>
F;1781-1863/Domain: fibronectin type III repeat homology <FN3N>
F;1970-1972/Region: cell attachment (R-G-D) motif
F;1982-2062/Domain: fibronectin type III repeat homology <FN3O>
F;1985-2216/Domain: fibrin binding <FB2>
F;2085-2124/Domain: fibronectin type I repeat homology <1F10>
F;2130-2167/Domain: fibronectin type I repeat homology <1F11>
F;2174-2209/Domain: fibronectin type I repeat homology <1F12>
F;3/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of fibrin) #status experimental
F;21-47,45-56,66-94,92-104,110-138,136-148,155-184,182-194,200-229,227-239,277-304,302-3
7,2155-2167,2174-2200,2198-2209/Disulfide bonds: #status predicted
F;399,497,511,846,976,1213,1987/Binding site: carbohydrate (Asn) (covalent) #status exper
F;1205,1692/Binding site: carbohydrate (Asn) (covalent) #status absent
F;1943,1944/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;2246/Disulfide bonds: interchain (to 2250) #status predicted
F;2250/Disulfide bonds: interchain (to 2246) #status predicted
F;2263/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 96.5%; Score 605; DB 1; Length 2265;
Best Local Similarity 96.4%; Pred. No. 1.9e-49;
Matches 106; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PIAKCFDHRAGTSYVVGTEWKPQGMWVDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
DB 150 PIAKCFDHRAGTSYVVGTEWKPQGMWVDCTCLGEGSGRITCTSRNRCNDQDRTSY 209
QY 61 RIGDTWSKKDNRGNLLQICITGNRGWKCERHTSVQTTSSGSGPFTDVR 110
DB 210 RIGDTWSKKDNRGNLLQICITGNRGWKCERHTSVQTTSSGSGPFTDVR 259

RESULT 3
S14428
fibronectin precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Oct-1994 #sequence revision 28-Oct-1994 #text change 09-Jul-2004
C;Accession: S14428; S12455; A22319; S46203; S00459; A27252; I59049
R;Hynes, R.O.
submitted to the EMBL Data Library, July 1989
A;Reference number: S14428
A;Accession: S14428
A;Molecule type: mRNA
A;Residues: 1-2477 <HYN>
A;Cross-references: UNIPROT:P04937; EMBL:X15906; NID:G56163; PIDN:CAA34020.1; PID:G5616
R;Schwarzbauer, J.E.; Patel, R.S.; Fonda, D.; Hynes, R.O.
EMBO J. 6, 2573-2580, 1987
A;Title: Multiple sites of alternative splicing of the rat fibronectin gene transcript.
A;Reference number: S12455; MUID:88054951; PMID:2445560
A;Accession: S12455
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 609-1810,'T',1812-2283 <SCH>
A;Cross-references: EMBL:X15906
R;Tamkun, J.W.; Schwarzbauer, J.E.; Hynes, R.O.
Proc. Natl. Acad. Sci. U.S.A. 81, 5140-5144, 1984
A;Title: A single rat fibronectin gene generates three different mRNAs by alternative sp
A;Reference number: A22319; MUID:84238097; PMID:6089177
A;Accession: A22319
A;Molecule type: DNA
A;Residues: 2052-2237 <TAM>
R;Falkenberg, C.; Enghild, J.J.; Thøgersen, I.B.; Salvesen, G.; Akerstroem, B.
Biochem. J. 301, 745-751, 1994

Db 241 VVTKHLYTKNGSSAGP 258

RESULT 12

Tl0756

Nel-homolog protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: Tl0756

R;Kuroda, S.; Tokunaga, C.; Kiyohara, Y.; Konishi, H.; Matsuhaishi, S.; Kikkawa, U.

submitted to the EMBL Data Library, November 1998

A;Description: Protein kinase C-binding protein.

A;Reference number: Z17122

A;Accession: Tl0756

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-810 <UNP>

A;Cross-references: UNIPROT:Q62919; EMBL:U48246; NID:g3851179; PID:g3851180

A;Experimental source: strain Sprague-Dawley, brain

Query Match 11.3%; Score 71; DB 2; Length 810;

Best Local Similarity 33.3%; Pred. No. 42;

Matches 20; Conservative 6; Mismatches 22; Indels 12; Gaps 3;

QY 31 VDCTCLGEGSGRITCTSRNRCNDQDTRTSYRIGDTWSKKNRGNLLQICITGNRGGEWKC 90

Db 678 VDLFCCECDTRVT---SQCLDQSGQKLYRSGDNWTH-----SCQQCRCL---EGEADC 725

RESULT 13

I46059

beta-1 integrin subunit - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004

C;Accession: I46059

R;MacLaren, L.A.; Wildeman, A.G.

Biol. Reprod. 53, 153-165, 1995

A;Title: Fibronectin receptors in preimplantation development: cloning, expression, and

A;Reference number: I46059; MUID:95399478; PMID:7545439

A;Accession: I46059

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-773 <MAC>

A;Cross-references: UNIPROT:P53712; EMBL:U10865; NID:g520520; PIDN:AAA80571.1; PID:g5205

C;Superfamily: integrin beta chain; laminin-type EGF-like homology

Query Match 11.2%; Score 70.5; DB 2; Length 773;

Best Local Similarity 24.0%; Pred. No. 45;

Matches 31; Conservative 15; Mismatches 44; Indels 39; Gaps 7;

QY 3 ABKCFDHAAGTSYV--VGTWEPYQGWMMVDTCLGEG-SGRITCTSRN-----RC 51

Db 409 ANKCPNKNSETIKIKPLGTFEEVILQICECEQGEIGPSPKCHDNGTFCGCACRC 468

QY 52 NDQDTRTSYRIG-----DTWSKKNRGNLLQICITGNR---GEWKCERH 93

Db 469 NEG-----RVGRHCCESTDVNSDMDAYCRKNSSE-----ICSNNGECVCGQCVCKR 518

QY 94 TSVQTTSSG 102

Db 519 DNTNIIYSG 527

RESULT 14

T04777

hypothetical protein F10M10.80 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C;Accession: T04777

R;Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Hoheisel, J.; Mewes, H.W.; Mayer,

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15384

A;Accession: T04777

A;Molecule type: DNA

A;Residues: 1-1074 <BEV>

A;Cross-references: UNIPROT:Q9SYZ6; EMBL:AL035521

A;Experimental source: cultivar Columbia; BAC clone F10M10

C;Genetics:

A;Map position: 4

A;Note: F10M10.80

C;Superfamily: Arabidopsis thaliana hypothetical protein F10M10.80

Query Match 11.2%; Score 70.5; DB 2; Length 1074;

Best Local Similarity 27.7%; Pred. No. 61;

Matches 31; Conservative 9; Mismatches 31; Indels 41; Gaps 8;

QY 19 ETWEKPYQGWMMVDTCLGEGSGRITCTSRNRCNDQDTRTSYRI-----G 63

Db 820 ETWCK----W-LDDC-----AKGNISC-----CNDPKTQSYARASLLNVCNQDQSGSG 864

QY 64 DTWSKKNRGNL-LQICITGN-----GRGEWKC-----ERHTSVQTTSSGS 104

Db 865 DGGSSKPDISNMNSNCPRYGDMIFLNPGLPHWKHEKERQSGKKNESSEG 916

RESULT 15

T18423

hypothetical protein C0150w - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T18423

R;Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, July 1997

A;Reference number: Z18934

A;Accession: T18423

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1169 <LAW>

A;Cross-references: UNIPROT:O77317; EMBL:Z97348; NID:e1323671; PIDN:CAB10579.1

C;Genetics:

A;Introns: 26/2; 1102/3; 1136/2

A;Note: C0150w

Query Match 11.2%; Score 70.5; DB 2; Length 1169;

Best Local Similarity 21.5%; Pred. No. 66;

Matches 17; Conservative 21; Mismatches 32; Indels 9; Gaps 2;

QY 1 PIAEKCFDHAAGTSYVVGTEWEPYQGWMMVDTCLGEGSGRITCTSRNRCNDQDTRTSY 60

Db 811 PLSILAVEKLLNSLTLSEVKDKSWRNW-----FGVSSTEYDNTTNNKSTNKTNTN 863

QY 61 RIGDTWSKKD--NRGNLLQ 77

Db 864 QIEDKXKKKQGGKSNIIIE 882

Search completed: November 3, 2004, 23:55:43

Job time : 12.3014 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:34:19 ; Search time 58.0137 Seconds
(without alignments)
1090.970 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAEKFCDHAAGTSYVVGET.....ERHTSVQTTSSGSGPFTDVR 110

Scoring table: BLOSUM62

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	100.0	749	2	Q7Z391
2	627	100.0	1103	2	Q6MZP4
3	627	100.0	1103	2	Cae45885
4	627	100.0	1103	2	Cae46200
5	627	100.0	2296	2	Q6N0A6
6	627	100.0	2296	2	Cae45714
7	627	100.0	2386	1	FINC HUMAN
8	627	100.0	2477	2	Q6MZU5
9	627	100.0	2477	2	Cae45932
10	623	99.4	2444	2	Q6N025
11	623	99.4	2444	2	Cae45847
12	605	96.5	2265	1	FINC BOVIN
13	581.5	92.7	2477	1	FINC MOUSE
14	580.5	92.6	2477	1	FINC RAT
15	551.5	88.0	296	2	Q8C6J7
16	539	86.0	2481	1	FINC XENLA
17	539	86.0	2481	1	Q6GQA5
18	520	82.9	922	2	Q93405
19	520	82.9	2478	2	O93406
20	464.5	74.1	2408	2	Q6JAN2
21	464.5	74.1	2408	2	AAT08488
22	224	35.7	522	1	FINC HORSE
23	223	35.6	351	2	O02816
24	218	34.8	522	1	FINC CANFA
25	216	34.4	320	2	Q95KV4
26	216	34.4	412	2	Q71U44
27	216	34.4	412	2	AAD33692
28	213	34.0	195	2	Q7T2W7
29	213	34.0	211	2	Q7L553
30	213	34.0	211	2	AH05858
31	213	34.0	268	2	Q6PJE5

32	213	34.0	268	2	AAH16875	Aah16875 homo sapi
33	213	34.0	358	2	Q7T1S2	Q7t1s2 brachydanio
34	213	34.0	1255	2	Q6MZS0	Q6mzso homo sapien
35	213	34.0	1255	2	Cae45958	Cae45958 homo sapi
36	213	34.0	2193	2	Q6MZM7	Q6mzm7 homo sapien
37	213	34.0	2193	2	Cae46002	Cae46002 homo sapi
38	211	33.7	216	2	Q99KD0	Q99kd0 mus musculu
39	211	33.7	810	2	Q8R3F3	Q8r3f3 mus musculu
40	206	32.9	1034	2	Q6N084	Q6n084 homo sapien
41	206	32.9	1034	2	Cae45786	Cae45786 homo sapi
42	196.5	31.3	1328	1	FINC PLEWA	Q91289 pleurodeles
43	192	30.6	1256	1	FINC_CHICK	P11722 gallus gall
44	182	29.0	141	2	Q90XQ2	Q90xq2 ambystoma m
45	173.5	27.7	190	1	FINC_NOTVI	Q91400 notophthalm

ALIGNMENTS

RESULT 1
Q7Z391 PRELIMINARY; PRT; 749 AA.
AC Q7Z391;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DE Hypothetical protein DKF2p686B18150.
GN Name=DKF2p686B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human colon endothelial primary cell culture;
RA Bloeker H., Boecker M., Mewes H.W., Weil B., Amlid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; BX538045; CAD97984.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 9.
DR PROSITE; PS00023; FIBRONECTIN 2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 627; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 3.3e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 60		PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db	273	PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 332		PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 332
Qy	61	RIGDTWSKDNKRNLLQICITGNRGKWCERHTSVQTTSSGSGPFTDVR 110		RIGDTWSKDNKRNLLQICITGNRGKWCERHTSVQTTSSGSGPFTDVR 110
Db	333	RIGDTWSKDNKRNLLQICITGNRGKWCERHTSVQTTSSGSGPFTDVR 362		RIGDTWSKDNKRNLLQICITGNRGKWCERHTSVQTTSSGSGPFTDVR 362

RESULT 2
Q6MZP4 PRELIMINARY; PRT; 1103 AA.
AC Q6MZP4;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)

05-JUL-2004 (Tremblrel. 27, Last annotation update)
Hypothetical protein DKFZp686K139 (Hypothetical protein DKFZp686F219)
(Fragment).
GN Name=DKFZp686K139; Synonyms=DKFZp686F219;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN
RP
SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RG THE GERMAN HUMAN CDNA CONSORTIUM;
RA Ansgorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE45885.1; -;
DR EMBL; BX640802; CAE45885.1; -;
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 4.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRODOM; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 4.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 4.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;
Query Match 100.0%; Score 627; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 5e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGTEWKPYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 270 PIAKCFDHAAGTSYVVGTEWKPYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 329
Qy 61 RIGDTSKKNRGNLLQICCTGNGRGEWKRCERTSVQTTSSGSGPFTDVR 110
Db 330 RIGDTSKKNRGNLLQICCTGNGRGEWKRCERTSVQTTSSGSGPFTDVR 379
RESULT 3
CAE45885
ID CAE45885 PRELIMINARY; PRT; 1103 AA.
AC CAE45885;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686F219 (Fragment).
GN DKFZp686F219.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN
RP
SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RA Ansgorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640802; CAE45885.1; -;
KW Hypothetical protein.

FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;
Query Match 100.0%; Score 627; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 5e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGTEWKPYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 270 PIAKCFDHAAGTSYVVGTEWKPYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 329
Qy 61 RIGDTSKKNRGNLLQICCTGNGRGEWKRCERTSVQTTSSGSGPFTDVR 110
Db 330 RIGDTSKKNRGNLLQICCTGNGRGEWKRCERTSVQTTSSGSGPFTDVR 379
RESULT 4
CAE46200
ID CAE46200 PRELIMINARY; PRT; 1103 AA.
AC CAE46200;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DT 02-MAR-2004 (Tremblrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686K139 (Fragment).
GN DKFZp686K139.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN
RP
SEQUENCE FROM N.A.
RC TISSUE=Human cervix;
RA Ansgorge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX649182; CAE46200.1; -;
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1103 AA; 122113 MW; 82FEC4CAF634AD56 CRC64;
Query Match 100.0%; Score 627; DB 2; Length 1103;
Best Local Similarity 100.0%; Pred. No. 5e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGTEWKPYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 270 PIAKCFDHAAGTSYVVGTEWKPYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 329
Qy 61 RIGDTSKKNRGNLLQICCTGNGRGEWKRCERTSVQTTSSGSGPFTDVR 110
Db 330 RIGDTSKKNRGNLLQICCTGNGRGEWKRCERTSVQTTSSGSGPFTDVR 379
RESULT 5
Q6NOA6
ID Q6NOA6 PRELIMINARY; PRT; 2296 AA.
AC Q6NOA6;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Hypothetical protein DKFZp686M04163.
GN Name=DKFZp686M04163;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
RN
RP
SEQUENCE FROM N.A.
RC TISSUE=Human uterus endothel primary cell culture;
RG THE GERMAN HUMAN CDNA CONSORTIUM;
RA Bloecker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.

```
DR EMBL; BX640608; CAB45714.1; --
DR InterPro; IPR002086; Aldehyde_dehydr.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; FIBRIN1.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000562; FN_Type_II.
DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 12.
DR SMART; SM00059; FN2; 2.
DR SMART; SM00060; FN3; 15.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
DR PROSITE; PS00822; EGF_1; UNKNOWN_2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS00853; FN3; 15.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 2296 AA; 252761 MW; 9AB2D723CC0CED70 CRC64;

Query Match 100.0%; Score 627; DB 2; Length 2296;
Best Local Similarity 100.0%; Pred. No. 1.1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 181 PIAKCFDHAAGTSYVVGTEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 240

Qy 61 RIGDTWSKKDNRGNLLQCICTGNGRGEWKCEKRTSVQTTSSSGGPTDVR 110
Db 241 RIGDTWSKKDNRGNLLQCICTGNGRGEWKCEKRTSVQTTSSSGGPTDVR 290

RESULT 6
CAB45714 PRELIMINARY; PRT; 2296 AA.
ID CAB45714
AC CAB45714;
DT 02-MAR-2004 (TRENBLrel. 27, Created)
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)
DE Hypothetical protein DKFZp686M04163.
GN DKFZP686M04163.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
TISSUE=Human uterus endothel primary cell culture;
RA Bloeker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640608; CAB45714.1; --
KW Hypothetical protein.
SQ SEQUENCE 2296 AA; 252761 MW; 9AB2D723CC0CED70 CRC64;

Query Match 100.0%; Score 627; DB 2; Length 2296;
Best Local Similarity 100.0%; Pred. No. 1.1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 181 PIAKCFDHAAGTSYVVGTEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 240

Qy 61 RIGDTWSKKDNRGNLLQCICTGNGRGEWKCEKRTSVQTTSSSGGPTDVR 110
Db 241 RIGDTWSKKDNRGNLLQCICTGNGRGEWKCEKRTSVQTTSSSGGPTDVR 290

RESULT 7
FINC_HUMAN STANDARD; PRT; 2386 AA.
ID P02751; Q95609; Q95610; Q14312; Q14325; Q14326; Q86T27; Q8IVIG;
AC Q96KP7; Q96KP8; Q96KP9; Q9H1B8; Q9HAP3; Q9UMK2;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Fibronectin precursor (FN) (Cold-insoluble globulin) (CIG).
GN Name=FN1; Synonyms=FN;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21600194; PubMed=11737888;
RA Schor S.L., Schor A.M.;
RT "Phenotypic and genetic alterations in mammary stroma: implications
RT for tumour progression.";
RL Breast Cancer Res. 3:373-379(2001).
RN [2]
SEQUENCE FROM N.A. (ISOFORMS 3; 7 AND 10).
RC TISSUE=Cervix;
RA Ansoerge W., Krieger S., Regiert T., Rittmueller C., Schwager B.,
RA Mewes H.-W., Weil B., Amid C., Osanger A., Fobo G., Han M.,
RA Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE OF 1-38 FROM N.A.
RX MEDLINE=87030890; PubMed=3770189;
RA Gutman A., Yamada K.M., Kornblitt A.R.;
RT "Human fibronectin is synthesized as a pre-propolypeptide.";
RL FEBS Lett. 207:145-148(1986).
RN [4]
SEQUENCE OF 1-49 FROM N.A.
RX MEDLINE=87175578; PubMed=3031656;
RA Dean D.C., Bowlus C.L., Bourgeois S.;
RT "Cloning and analysis of the promoter region of the human fibronectin
RT gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:1876-1880(1987).
RN [5]
SEQUENCE OF 28-2386 FROM N.A. (ISOFORM 3).
RX MEDLINE=85284965; PubMed=2992939;
RA Kornblitt A.R., Umezawa K., Vibe-Pedersen K., Baralle F.E.;
RT "Primary structure of human fibronectin: differential splicing may
RT generate at least 10 polypeptides from a single gene.";
RL EMBO J. 4:1755-1759(1985).
RN [6]
SEQUENCE OF 103-481 AND 2228-2386 FROM N.A. (ISOFORMS 1; 3; 8 AND 9).
RC TISSUE=Periphereal blood T-cell, and Umbilical vein endothelial cells;
RA Godfrey H.P., Ebrahim A.A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [7]
SEQUENCE OF 973-2386 FROM N.A. (ISOFORM 3).
RX MEDLINE=84272258; PubMed=6462919;
RA Kornblitt A.R., Vibe-Pedersen K., Baralle F.E.;
RT "Human fibronectin: cell specific alternative mRNA splicing generates
RT polypeptide chains differing in the number of internal repeats.";
RL Nucleic Acids Res. 12:5853-5868(1984).
RN [8]
SEQUENCE OF 1232-1782 FROM N.A. (ISOFORM 7).
RX MEDLINE=88233940; PubMed=3375063;
RA Paolella G., Henschliffe C., Sebastio G., Baralle F.E.;
RT "Sequence analysis and in vivo expression show that alternative
RT splicing of ED-B and ED-A regions of the human fibronectin gene are
RT independent events.";
RL Nucleic Acids Res. 16:3545-3557(1988).
RN [9]
SEQUENCE OF 1257-1365 FROM N.A. (ISOFORM 11).
RX MEDLINE=88041070; PubMed=3478690;
RA Gutman A., Kornblitt A.R.;
RT "Identification of a third region of cell-specific alternative
```

RT splicing in human fibronectin mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7179-7182(1987).
RN [10]
RP SEQUENCE OF 1441-1548.
RX MEDLINE=82265604; PubMed=7050098;
RA Pierschbacher M.D., Ruoslahti E., Sundelin J., Lind P., Peterson P.A.;
RT "The cell attachment domain of fibronectin. Determination of the
RT primary structure.";
RL J. Biol. Chem. 257:9593-9597(1982).
RN [11]
RP SEQUENCE OF 1448-1540 FROM N.A.
RX MEDLINE=83290929; PubMed=6688418;
RA Oldberg A., Linney E., Ruoslahti E.;
RT "Molecular cloning and nucleotide sequence of a cDNA clone coding for
RT the cell attachment domain in human fibronectin.";
RL J. Biol. Chem. 258:10193-10196(1983).
RN [12]
RP SEQUENCE OF 1448-1540 FROM N.A.
RX MEDLINE=86111901; PubMed=3003095;
RA Oldberg A., Ruoslahti E.;
RT "Evolution of the fibronectin gene. Exon structure of cell attachment
RT domain.";
RL J. Biol. Chem. 261:2113-2116(1986).
RN [13]
RP SEQUENCE OF 1594-2386 FROM N.A. (ISOFORM 1).
RX MEDLINE=85280409; PubMed=2992573;
RA Bernard M.P., Kolbe M., Weil D., Chu M.-L.;
RT "Human cellular fibronectin: comparison of the carboxyl-terminal
RT portion with rat identifies primary structural domains separated by
RT hypervariable regions.";
RL Biochemistry 24:2698-2704(1985).
RN [14]
RP SEQUENCE OF 1712-1739 FROM N.A.
RX MEDLINE=87026578; PubMed=3021206;
RA Sekiguchi K., Klos A.M., Kurachi K., Yoshitake S., Hakomori S.;
RT "Human liver fibronectin complementary DNAs: identification of two
RT different messenger RNAs possibly encoding the alpha and beta subunits
RT of plasma fibronectin.";
RL Biochemistry 25:4936-4941(1986).
RN [15]
RP SEQUENCE OF 1788-2386 FROM N.A. (ISOFORMS 4; 5 AND 6).
RC TISSUE=Cartilage;
RX MEDLINE=22126816; PubMed=12127932;
RA Parker A.E., Boutell J., Carr A., Maciewicz R.A.;
RT "Novel cartilage-specific splice variants of fibronectin.";
RL Osteoarthritis Cartilage 10:528-534(2002).
RN [16]
RP SEQUENCE OF 32-290.
RX MEDLINE=84032463; PubMed=6630202;
RA Garcia-Pardo A., Pearlsstein E., Frangione B.;
RT "Primary structure of human plasma fibronectin. The 29,000-dalton NH2-
RT terminal domain.";
RL J. Biol. Chem. 258:12670-12674(1983).
RN [17]
RP SEQUENCE OF 309-608, AND COLLAGEN-BINDING.
RX MEDLINE=87080265; PubMed=3024962;
RA Owens R.J., Baralle F.E.;
RT "Mapping the collagen-binding site of human fibronectin by expression
RT in *Escherichia coli*.";
RL EMBO J. 5:2825-2830(1986).
RN [18]
RP SULFATION.
RX MEDLINE=86042625; PubMed=2414772;
RA Liu M.C., Yu S., Sy J., Redman C.M., Lipmann F.;
RT "Tyrosine sulfation of proteins from the human hepatoma cell line
RT HepG2.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7160-7164(1985).
RN [19]
RP O-GLYCOSYLATION OF THR-2064.
RX MEDLINE=91190085; PubMed=2012601;
RA Tressel T., McCarthy J.B., Calaycay J., Lee T.D., Legesse K.,
RA Shively J.E., Pande H.;
RT "Human plasma fibronectin. Demonstration of structural differences

RT between the A- and B-chains in the III CS region.";
RL Biochem. J. 274:731-738(1991).
RN [20]
RP FBLN1-BINDING SITE.
RX MEDLINE=93015879; PubMed=1400330;
RA Balbona K., Tran H., Godyna S., Strickland D.K.,
RA Argaves W.S.;
RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding
RT region of fibronectin.";
RL J. Biol. Chem. 267:20120-20125(1992).
RN [21]
RP CHARACTERIZATION OF FIBRIN-BINDING SITE 1.
RX MEDLINE=95081153; PubMed=7989369;
RA Rostagno A., Williams M.J., Baron M., Campbell I.D., Gold L.I.;
RT "Further characterization of the NH2-terminal fibrin-binding site on
RT fibronectin.";
RL J. Biol. Chem. 269:31938-31945(1994).
RN [22]
RP INTERACTION WITH LGALS3BP.
RX PubMed=9501082;
RA Sasaki T., Brakebusch C., Engel J., Timpl R.;
RT "Mac-2 binding protein is a cell-adhesive protein of the extracellular
RT matrix which self-assembles into ring-like structures and binds beta1
RT integrins, collagens and fibronectin.";
RL EMBO J. 17:1606-1613(1998).
RN [23]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE=92162710; PubMed=1311202;
RA Baron M., Main A.L., Driscoll P.C., Mardon H.J., Boyd J.,
RA Campbell I.D.;
RT "1H NMR assignment and secondary structure of the cell adhesion type
RT III module of fibronectin.";
RL Biochemistry 31:2068-2073(1992).
RN [24]
RP STRUCTURE BY NMR OF 1447-1540.
RX MEDLINE=93046665; PubMed=1423622;
RA Main A.L., Harvey T.S., Baron M., Boyd J., Campbell I.D.;
RT "The three-dimensional structure of the tenth type III module of
RT fibronectin: an insight into RGD-mediated interactions.";
RL Cell 71:671-678(1992).
RN [25]
RP STRUCTURE BY NMR OF 182-275.
RX MEDLINE=94141923; PubMed=8308892;
RA Williams M.J., Phan I., Harvey T.S., Rostagno A., Gold L.I.,
RA Campbell I.D.;
RT "Solution structure of a pair of fibronectin type 1 modules with
RT fibrin binding activity.";
RL J. Mol. Biol. 235:1302-1311(1994).
RN [26]
RP STRUCTURE BY NMR OF 32-92.
RX MEDLINE=96069779; PubMed=7583666;
RA Potts J.R., Phan I., Williams M.J., Campbell I.D.;
RT "High-resolution structural studies of the factor XIIIa crosslinking
RT site and the first type 1 module of fibronectin.";
RL Nat. Struct. Biol. 2:946-950(1995).
RN [27]
RP Query Match 100.0%; Score 627; DB 1; Length 2386;
Best Local Similarity 100.0%; Pred. No. 1.1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PIAKCFDHAAGTSYVVGTEWPKYQGMWVDTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 181 PIAKCFDHAAGTSYVVGTEWPKYQGMWVDTCLGEGSGRITCTSRNRCNDQDTRTSY 240
QY 61 RIGDTWSKKDNRGNLLQCICTGNRGCEWKERTHSVQTTSSSGSPFTDVR 110
Db 241 RIGDTWSKKDNRGNLLQCICTGNRGCEWKERTHSVQTTSSSGSPFTDVR 290
RESULT 8
Q6MZU5 PRELIMINARY; PRT; 2477 AA.
ID Q6MZU5


```

RESULT 11
CAE45847 ID CAE45847 PRELIMINARY; PRT; 2444 AA.
AC CAE45847;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein DKFP686M2451 (Fragment).
GN DKFP686M2451.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human endometrium carcinoma cell line;
RA Poustka A., Albert R., Moosmayer P., Schnupp I., Wellenreuther R.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX640731; CAE45847.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 2444 AA; 268676 MW; 71C5E8C56A84C7BC CRC64;

Query Match 99.4%; Score 623; DB 2; Length 2444;
Best Local Similarity 99.1%; Pred. No. 3e-56;
Matches 109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGTEWKPQYQGMWVDCICLGGSGRITCTSRNRCNDQDTRTSY 60
DQ PIAKCFDHAAGTSYVVGTEWKPQYQGMWVDCICLGGSGRITCTSRNRCNDQDTRTSY 329
DQ 61 RIGTWSKKNRGNLLQICITGNGRGEWKGCERHTSVQTTSSGSGPFTDVR 110
DQ 330 RIGTWSKKNRGNLLQICITGNGRGEWKGCERHTSVQTTSSGSGPFTDAR 379

RESULT 12
FINC_BOVIN ID FINC_BOVIN STANDARD; PRT; 2265 AA.
AC P07589;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Fibronectin (FN).
GN Name=FN1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=87054047; PubMed=3780752;
RA Skorstengaard K., Jensen M.S., Sahl P., Petersen T.E., Magnusson S.;
RT "Complete primary structure of bovine plasma fibronectin.";
RL Eur. J. Biochem. 161:441-453(1986).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=83117805; PubMed=6218503;
RA Petersen T.E., Thorgeresen H.C., Skorstengaard K., Vibe-Pedersen K.,
RA Sahl P., Sottrup-Jensen L., Magnusson S.;
RT "Partial primary structure of bovine plasma fibronectin: three types
of internal homology.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:137-141(1983).
RN [3]
RP SEQUENCE OF 2170-2265 FROM N.A.
RX MEDLINE=83221567; PubMed=6304699;
RA Kornblihtt A.R., Vibe-Pedersen K., Baralle F.E.;
RT "Isolation and characterization of cDNA clones for human and bovine
fibronectins.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3218-3222(1983).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
including collagen, fibrin, heparin, DNA, and actin. Fibronectins
are involved in cell adhesion, cell motility, opsonization, wound
healing, and maintenance of cell shape.
-!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
variants, connected by 2 disulfide bonds near the carboxyl ends;
to a lesser extent homodimers.
-!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=1;
Comment=A number of isoforms are produced. Each of the "extra
domain" and the connecting strand 3 are present in some forms of
fibronectin and absent in others;
Name=1;
IsoID=P07589-1; Sequence=Displayed;
-!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
forms), made by fibroblasts, epithelial and other cell types, is
deposited as fibrils in the extracellular matrix.
-!- PTM: Sulfated (By similarity).
-!- SIMILARITY: Contains 12 fibronectin type I domains.
-!- SIMILARITY: Contains 2 fibronectin type II domains.
-!- SIMILARITY: Contains 15 fibronectin type III domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; K00800; AAA30521.2; -
PIR; A26452; FNEO.
HSP; P08253; IKS0.
InterPro; IPR006209; EGF_like.
InterPro; IPR000083; Fibrctnl.
InterPro; IPR003962; FNIII subd.
InterPro; IPR003961; FN_III.
InterPro; IPR008957; FN_III-like.
InterPro; IPR000562; FN_Type_II.
Pfam; PF00039; fnl; 12.
Pfam; PF00040; fn2; 2.
Pfam; PF00041; fn3; 15.
PRINTS; PR00012; FNTYPEI.
PRINTS; PR00013; FNTYPEII.
PRINTS; PR00014; FNTYPEIII.
ProDom; PD000995; FN_Type_II; 2.
PROSITE; PS00022; EGF 1; 2.
PROSITE; PS01253; FIBRONECTIN_1; 12.
PROSITE; PS00023; FIBRONECTIN_2; 2.
PROSITE; PS50853; FN3; 15.
KW Acute phase; Alternative splicing; Cell adhesion; Cell shape;
KW Direct protein sequencing; Glycoprotein; Heparin-binding;
KW Phosphorylation; Plasma; Pyrrolidone carboxylic acid; Repeat;
Sulfation.
FT MOD_RES 1 1 Pyrrolidone carboxylic acid.
FT DOMAIN 21 241 Fibrin- and heparin-binding 1.
FT DOMAIN 277 577 Collagen-binding.
FT DNA_BIND 876 1141 CELL-ATTACHMENT.
FT DOMAIN 1236 1509 Heparin-binding 2.
FT DOMAIN 1600 1870 Fibrin-binding 2.
FT DOMAIN 1991 2216 Fibronectin type-I 1.
FT DOMAIN 19 59 Fibronectin type-I 2.
FT DOMAIN 64 107 Fibronectin type-I 3.
FT DOMAIN 108 151 Fibronectin type-I 4.
FT DOMAIN 153 197 Fibronectin type-I 5.
FT DOMAIN 198 242 Fibronectin type-I 6.
FT DOMAIN 275 314 Fibronectin type-II 1.
FT DOMAIN 314 373 Fibronectin type-II 2.
FT DOMAIN 374 438 Fibronectin type-I 7.
FT DOMAIN 437 480 Fibronectin type-I 8.
FT DOMAIN 485 527 Fibronectin type-I 9.
FT DOMAIN 528 571

```

FT	DOMAIN	576	668	Fibronectin type-III 1.	DB	210	RIGDTWSKDNKRNLLQICITGNGRGEWKCRHRTSLQTTSGSGFTDVR	259
FT	DOMAIN	689	778	Fibronectin type-III 2.				
FT	DOMAIN	780	867	Fibronectin type-III 3.				
FT	DOMAIN	877	964	Fibronectin type-III 4.				
FT	DOMAIN	965	1053	Fibronectin type-III 5.				
FT	DOMAIN	1056	1141	Fibronectin type-III 6.				
FT	DOMAIN	1142	1234	Fibronectin type-III 7.				
FT	DOMAIN	1235	1325	Fibronectin type-III 8.				
FT	DOMAIN	1326	1415	Fibronectin type-III 9.				
FT	DOMAIN	1416	1505	Fibronectin type-III 10.				
FT	DOMAIN	1510	1599	Fibronectin type-III 11 (extra domain).				
FT	DOMAIN	1602	1689	Fibronectin type-III 12.				
FT	DOMAIN	1692	1780	Fibronectin type-III 13.				
FT	DOMAIN	1781	1870	Fibronectin type-III 14.				
FT	DOMAIN	1871	1990	CONNECTING STRAND 3 (CS-3) (V REGION).				
FT	DOMAIN	1979	2069	Fibronectin type-III 15.				
FT	DOMAIN	2083	2127	Fibronectin type-I 10.				
FT	DOMAIN	2128	2170	Fibronectin type-I 11.				
FT	DOMAIN	2172	2215	Fibronectin type-I 12.				
FT	SITE	1493	1495	Cell attachment site.				
FT	DISULFID	21	47					
FT	DISULFID	45	56					
FT	DISULFID	66	94					
FT	DISULFID	92	104					
FT	DISULFID	110	138					
FT	DISULFID	136	148					
FT	DISULFID	155	184					
FT	DISULFID	182	194					
FT	DISULFID	200	229					
FT	DISULFID	227	239					
FT	DISULFID	277	304					
FT	DISULFID	302	311					
FT	DISULFID	329	355					
FT	DISULFID	343	370					
FT	DISULFID	389	415					
FT	DISULFID	403	430					
FT	DISULFID	439	467					
FT	DISULFID	465	477					
FT	DISULFID	487	514					
FT	DISULFID	512	524					
FT	DISULFID	530	558					
FT	DISULFID	556	568					
FT	DISULFID	2085	2114					
FT	DISULFID	2112	2124					
FT	DISULFID	2130	2157					
FT	DISULFID	2155	2167					
FT	DISULFID	2174	2200					
FT	DISULFID	2198	2209					
FT	DISULFID	2246	2246					
FT	DISULFID	2250	2250					
FT	MOD RES	845	845					
FT	MOD RES	850	850					
FT	CARBOHYD	399	399					
FT	CARBOHYD	497	497					
FT	CARBOHYD	511	511					
FT	CARBOHYD	846	846					
FT	CARBOHYD	976	976					
FT	CARBOHYD	1213	1213					
FT	CARBOHYD	1987	1987					
FT	CARBOHYD	1943	1943					
FT	CARBOHYD	1944	1944					
FT	MOD RES	2263	2263	Phosphoserine.				
SQ	SEQUENCE	2265	2265	AA; 249557 MW; C2D21D486F498D5C CRC64;				
Query Match			96.5%;	Score 605; DB 1; Length 2265;				
Best Local Similarity			96.4%;	Pred. No. 2.2e-54;				
Matches 106; Conservative			2; Mismatches	2; Indels	0; Gaps	0;		
Qy	1	PIAEKCFDHAAGTSVVGCTWEKPYQGMWV	DTCLGSGSGRITCTSRNRCNDQDTRSY	60				
Db	150	PIAEKCFDQAAGTSVVGCTWEKPYQGMWV	DTCLGSGSGRITCTSRNRCNDQDTRSY	209				
Qy	61	RIGDTWSKDNKRNLLQICITGNGRGEWKCRHRTSVQTTSGSGFTDVR	110					


```

infected mouse kidney cells." ;
Exp. Cell Res. 202:464-470(1992).
[7]
STRUCTURE BY NMR OF 1447-1630.
MEDLINE=98202578; PubMed=9533887;
RA Copie V., Tomita Y., Akiyama S.K., Aota S., Yamada K.M., Venable R.M.,
RA Pastor R.W., Krueger S., Torchia D.A.;
RA "Solution structure and dynamics of linked cell attachment modules of
RT mouse fibronectin containing the RGD and synergy regions: comparison
RT with the human fibronectin crystal structure." ;
J. Mol. Biol. 277:663-682(1998).
[8]
DOWN-REGULATION BY GLUCOCORTICOID.
RP MEDLINE=21600963; PubMed=11737251;
RX Gu Y.-C., Tait J.F., Gullberg D., Timpl R., Ekblom M.;
RA "Glucocorticoids down-regulate the extracellular matrix proteins
RT fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma." ;
RL Eur. J. Haematol. 67:176-184(2001).
CC -!- FUNCTION: Fibronectins bind cell surfaces and various compounds
CC including collagen, fibrin, heparin, DNA, and actin. Fibronectins
CC are involved in cell adhesion, cell motility, opsonization, wound
CC healing, and maintenance of cell shape.
CC -!- SUBUNIT: Mostly heterodimers or multimers of alternatively spliced
CC variants, connected by 2 disulfide bonds near the carboxyl ends;
CC to a lesser extend homodimers. Interacts with FBLN1 and LGALS3BP
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=A number of isoforms are produced. Each of the "extra
CC domain" and the connecting strand 3 are present in some forms of
CC fibronectin and absent in others;
CC Name=1;
CC -!- TISSUE SPECIFICITY: Plasma FN (soluble dimeric form) is secreted
CC by hepatocytes. Cellular FN (dimeric or cross-linked multimeric
CC forms), made by fibroblasts, epithelial and other cell types, is
CC deposited as fibrils in the extracellular matrix.
CC -!- INDUCTION: Glucocorticoids suppressed mRNA expression and protein
CC synthesis.
CC -!- PTM: Sulfated (By similarity).
CC -!- SIMILARITY: Contains 12 fibronectin type I domains.
CC -!- SIMILARITY: Contains 2 fibronectin type II domains.
CC -!- SIMILARITY: Contains 17 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; BC051082; AAH51082.1; -
CC EMBL; Z22729; CAA80422.1; -
CC EMBL; X82402; CAA57796.1; -
CC EMBL; X93167; CAA63654.1; -
CC EMBL; M18194; AAA37636.1; -
CC EMBL; S45680; AAB23491.1; -
CC PIR; A49173; A49173.
CC PIR; I48349; I48349.
CC PDB; 1MNF; @=1446-1630.
CC PDB; 2MNF; @=1446-1630.
CC MGI; 95566; Fnl.
CC GO; GO:0007155; P:cell adhesion; IDA.
CC GO; GO:0007044; P:cell-substrate junction assembly; IDA.
CC GO; GO:0042060; P:wound healing; IMP.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR003962; FnlII subd.
CC InterPro; IPR003961; FnlIII.
CC InterPro; IPR008957; FN-III-like.
CC InterPro; IPR000562; FN_Type-II.

```

```

DR Pfam; PF00039; fn1; 12.
DR Pfam; PF00040; fn2; 2.
DR Pfam; PF00041; fn3; 17.
DR PRINTS; PR00014; FNYPEIII.
DR PRODOM; PD000995; FN_Type_II; 2.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01253; FIBRONECTIN_1; 12.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
DR PROSITE; PS05083; FN3; 17.
DR 3D-structure; Acute phase; Alternative splicing; Cell adhesion;
KW Glycoprotein; Heparin-binding; Phosphorylation; Plasma; Repeat;
KW Signal; Sulfation.
FT SIGNAL 1 32 By similarity.
FT CHAIN 33 2477 Fibrinectin.
FT DOMAIN 53 273 Fibrin- and heparin-binding 1.
FT DOMAIN 308 608 Collagen-binding.
FT DNA_BIND 906 1171
FT DOMAIN 1357 1630
FT DOMAIN 1811 2081
FT DOMAIN 2296 2427
FT DOMAIN 51 96
FT DOMAIN 96 140
FT DOMAIN 140 185
FT DOMAIN 185 230
FT DOMAIN 230 272
FT DOMAIN 306 343
FT DOMAIN 345 404
FT DOMAIN 405 469
FT DOMAIN 468 516
FT DOMAIN 516 559
FT DOMAIN 559 602
FT DOMAIN 607 699
FT DOMAIN 719 808
FT DOMAIN 810 897
FT DOMAIN 905 994
FT DOMAIN 995 1083
FT DOMAIN 1091 1171
FT DOMAIN 1172 1264
FT DOMAIN 1265 1355
FT DOMAIN 1356 1446
FT DOMAIN 1447 1536
FT DOMAIN 1537 1626
FT DOMAIN 1631 1720
FT DOMAIN 1721 1810
FT DOMAIN 1813 1900
FT DOMAIN 1903 1991
FT DOMAIN 1992 2081
FT DOMAIN 2082 2201
FT DOMAIN 2190 2280
FT DOMAIN 2294 2338
FT DOMAIN 2339 2381
FT DOMAIN 2383 2426
FT SITE 1614 1616 Cell attachment site.
FT SITE 2181 2183 Cell attachment site.
FT DISULFID 53 79 By similarity.
FT DISULFID 77 88 By similarity.
FT DISULFID 98 126 By similarity.
FT DISULFID 124 136 By similarity.
FT DISULFID 142 170 By similarity.
FT DISULFID 168 180 By similarity.
Query Match 92.7%; Score 581.5; DB 1; Length 2477;
Best Local Similarity 91.8%; Pred. NO. 7e-52; Indels 1; Gaps 1;
Matches 101; Conservative 6; Mismatches 2;
QY 1 PIAKCFDHAAGTSYVVGTEWPKYQWMMVDCITCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 182 PIAKCFDHAAGTSYVVGTEWPKYQWMMVDCITCLGEGSGRITCTSRNRCNDQDRTSY 241
QY 61 RIGDTWSKDNKRNGLLQCTCTGNRGCEWKCERHTSVQTTSSGSGPFTDVR 110
Db 242 RIGDTWSKDNKRNGLLQCTCTGNRGCEWKCERH-ALQSASAGSGSFTDVR 290

```



```
FT DOMAIN 2383 2426 Fibrinectin type-I 12.
FT SITE 1614 1616 Cell attachment site.
FT SITE 2181 2183 Cell attachment site.
FT DISULFID 53 79 By similarity.
FT DISULFID 77 88 By similarity.
FT DISULFID 98 126 By similarity.
FT DISULFID 124 136 By similarity.
FT DISULFID 142 170 By similarity.
FT DISULFID 168 180 By similarity.
FT DISULFID 187 216 By similarity.
FT DISULFID 214 226 By similarity.
FT DISULFID 232 261 By similarity.
FT DISULFID 259 271 By similarity.
FT DISULFID 308 335 By similarity.
FT DISULFID 333 342 By similarity.
FT DISULFID 360 386 By similarity.
FT DISULFID 374 401 By similarity.
FT DISULFID 420 446 By similarity.
FT DISULFID 434 461 By similarity.
FT DISULFID 470 498 By similarity.
FT DISULFID 496 508 By similarity.
FT DISULFID 518 545 By similarity.
FT DISULFID 543 555 By similarity.
FT DISULFID 561 589 By similarity.
FT DISULFID 587 599 By similarity.
FT DISULFID 2296 2325 By similarity.
FT DISULFID 2323 2335 By similarity.
FT DISULFID 2341 2368 By similarity.
FT DISULFID 2366 2378 By similarity.
FT DISULFID 2385 2409 By similarity.
FT DISULFID 2407 2423 By similarity.
FT DISULFID 2458 2458 Interchain (with C-2452).
FT DISULFID 2462 2462 Interchain (with C-2458).
FT MOD RES 875 875 Sulfotyrosine (Potential).
FT MOD RES 880 880 Sulfotyrosine (Potential).
FT MOD RES 2392 2392 Sulfotyrosine (Potential).
FT MOD RES 2475 2475 Phosphoserine (By similarity).
FT CARBOHYD 430 430 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 528 528 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 542 542 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 876 876 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1006 1006 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1243 1243 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 2198 2198 O-linked (GlcNAc...) (By similarity).
FT CARBOHYD 2154 2154 O-linked (GalNAc...) (By similarity).
FT VARSPLIC 1720 1809 Missing (in isoform 2).
FT VARSPLIC 2082 2106 Missing (in isoform 3).
FT VARSPLIC 2082 2200 Missing (in isoform 4).
FT CONFLICT 2318 2318 G -> A (in Ref. 3).
SQ SEQUENCE 2477 AA; 272510 MW; B4391A472ECEDEB5 CRC64;

Query Match 92.6%; Score 580.5; DB 1; Length 2477;
Best Local Similarity 91.8%; Pred. No. 8.9e-52;
Matches 101; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

Qy 1 PIAKCFDHAAGTSYVVGETWKPQGVWVMDCTCLGEGSGRITCTSRNCNDQDRTSY 60
Db 182 PIAKCFDHAAGTSYVVGETWKPQGVWVMDCTCLGEGSGRITCTSRNCNDQDRTSY 241

Qy 61 RIGTWSKKNRGNLLQICITCNGRGWKCERHTSVQTTSSGSGPTDVR 110
Db 242 RIGTWSKKNRGNLLQICITCNGRGWKCERHV-LQSASAGSGSFTDVR 290

RESULT 15
Q8C6J7 PRELIMINARY; PRT; 296 AA.
ID AC Q8C6J7;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
```

DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330027I09 product:fibrinectin 1, full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA the FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh K., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T., Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK054456; BAC35784.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctn1.
DR Pfam; PF00039; fnl; 5.

DR PRINTS; PRO0012; FNTYPEI.
DR SMART; SMO0058; FN1; 5.
DR PROSITE; PS00022; EGF 1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 5.
SQ SEQUENCE 296 AA; 32617 MW; A0FF5F4809FB6439 CRC64;

Query Match 88.0%; Score 551.5; DB 2; Length 296;
Best Local Similarity 87.5%; Pred. No. 1.1e-49;
Matches 98; Conservative 6; Mismatches 5; Indels 3; Gaps 2;

Qy 1 PIAEKCFDHAAGTSYVVGETWEKPYQGMMVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 182 PIAEKCFDHAAGTSYVVGETWEKPYQGMMVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 241

Qy 61 RIGDTWSKKDNRGNLLQCVCTGNGRGEWKCEKHTSVQTTSSGS--GPFTDVR 110
Db 242 RIGDTWSKKDNRGNLLQCVCTGNGRGEWKCEKHTSVQTTSSGS--GPFTDVR 292

Search completed: November 3, 2004, 23:54:38
Job time : 59.0137 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:42:25 ; Search time 13.3733 Seconds
(without alignments)
545.489 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627
Sequence: 1 PIAKCFDHAAGTSYVVGET.....ERHVSQTTSSGSGPFTDVR 110

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	100.0	2231	1	US-08-153-799-16 Sequence 16, Appl
2	627	100.0	2324	1	US-08-283-857-1 Sequence 1, Appl
3	627	100.0	2324	5	PCT-US95-09819-1 Sequence 1, Appl
4	627	100.0	2386	2	US-09-016-366A-12 Sequence 12, Appl
5	627	100.0	2386	4	US-09-961-403-1 Sequence 1, Appl
6	627	100.0	2446	2	US-08-551-356-2 Sequence 2, Appl
7	627	100.0	2446	5	PCT-US93-12687-2 Patent No. 5455158
8	622	99.2	2327	6	5455158-1 Sequence 14, Appl
9	213	34.0	188	1	US-08-142-49B-14 Sequence 12, Appl
10	213	34.0	422	2	US-08-836-854-12 Sequence 9, Appl
11	213	34.0	474	2	US-08-836-854-9 Sequence 20, Appl
12	127	20.3	48	2	US-08-982-597A-20 Sequence 20, Appl
13	127	20.3	48	3	US-09-136-218-20 Sequence 20, Appl
14	110.5	17.6	48	2	US-08-982-597A-22 Sequence 22, Appl
15	110.5	17.6	48	3	US-09-136-218-22 Sequence 22, Appl
16	106	16.9	43	2	US-08-982-597A-21 Sequence 21, Appl
17	106	16.9	43	3	US-09-136-218-21 Sequence 21, Appl
18	97.5	15.6	2476	2	US-08-276-967-2 Sequence 2, Appl
19	88	14.0	44	1	US-08-340-812-6 Sequence 6, Appl
20	88	14.0	44	1	US-08-459-064B-6 Sequence 6, Appl
21	88	14.0	44	2	US-08-460-421A-6 Sequence 6, Appl
22	88	14.0	44	5	PCT-US93-00909-6 Sequence 6, Appl
23	74.5	11.9	133	4	US-09-252-991A-20369 Sequence 20369, A
24	70.5	11.2	167	4	US-09-252-991A-31970 Sequence 31970, A
25	70	11.2	564	4	US-09-489-039A-12729 Sequence 12729, A
26	70	11.2	788	2	US-08-918-914-4 Sequence 4, Appl
27	70	11.2	816	2	US-08-820-170A-37 Sequence 37, Appl

28	70	11.2	816	3	US-09-055-699-37 Sequence 37, Appl
29	70	11.2	816	3	US-09-273-565-37 Sequence 37, Appl
30	70	11.2	816	3	US-09-565-538-37 Sequence 37, Appl
31	70	11.2	816	3	US-09-661-468-37 Sequence 37, Appl
32	70	11.2	816	4	US-09-976-165-37 Sequence 37, Appl
33	70	11.2	1036	4	US-10-140-002-142 Sequence 142, App
34	69.5	11.1	129	4	US-09-270-767-40390 Sequence 40390, A
35	69.5	11.1	129	4	US-09-270-767-55606 Sequence 55606, A
36	69.5	11.1	679	1	US-07-998-973A-15 Sequence 15, Appl
37	69.5	11.1	679	2	US-08-452-800-15 Sequence 15, Appl
38	69.5	11.1	679	4	US-09-636-215-776 Sequence 776, App
39	69.5	11.1	679	4	US-09-685-166A-776 Sequence 776, App
40	69.5	11.1	679	4	US-09-679-426-776 Sequence 776, App
41	69.5	11.1	679	5	PCT-US92-11353-15 Sequence 15, Appl
42	69.5	11.1	950	4	US-10-009-332-1 Sequence 1, Appl
43	68.5	10.9	608	4	US-09-130-491-13 Sequence 13, Appl
44	68.5	10.9	949	4	US-09-568-559-2 Sequence 2, Appl
45	68.5	10.9	967	4	US-09-130-491-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-153-799-16
; Sequence 16, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Protein
LOCATION: 1..2231
OTHER INFORMATION: /note= "Human fibronectin"
US-08-153-799-16

Query Match 100.0%; Score 627; DB 1; Length 2231;
Best Local Similarity 100.0%; Pred. No. 2.2e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKCFDHAAGTSYVVGETWKEPYQGMWVDCICLGEGRITCTSRNRCNDQDRTSY 60
Db 150 PIAEKCFDHAAGTSYVVGETWKEPYQGMWVDCICLGEGRITCTSRNRCNDQDRTSY 209

Qy 61 RIGDTSKDKNRGNLLQICICTGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 210 RIGDTSKDKNRGNLLQICICTGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 259

RESULT 2
US-08-283-857-1
; Sequence 1, Application US/08283857
; Patent No. 5792742
; GENERAL INFORMATION:
; APPLICANT: GOLD, Leslie I.
; APPLICANT: ROSTAGNO, Agueda A.
; APPLICANT: BARON, Martin
; APPLICANT: CAMPBELL, Iain D.
; APPLICANT: WILLIAMS, Michael, J.
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USERS THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,857
; FILING DATE: 01-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/714,134
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-283-857-1

Query Match 100.0%; Score 627; DB 1; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2.3e-59;

Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKCFDHAAGTSYVVGETWKEPYQGMWVDCICLGEGRITCTSRNRCNDQDRTSY 60
Db 150 PIAEKCFDHAAGTSYVVGETWKEPYQGMWVDCICLGEGRITCTSRNRCNDQDRTSY 209

Qy 61 RIGDTSKDKNRGNLLQICICTGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 210 RIGDTSKDKNRGNLLQICICTGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 259

RESULT 3
PCT-US95-09819-1
; Sequence 1, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD=1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2324 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-09819-1

Query Match 100.0%; Score 627; DB 5; Length 2324;
Best Local Similarity 100.0%; Pred. No. 2.3e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKCFDHAAGTSYVVGETWKEPYQGMWVDCICLGEGRITCTSRNRCNDQDRTSY 60
Db 150 PIAEKCFDHAAGTSYVVGETWKEPYQGMWVDCICLGEGRITCTSRNRCNDQDRTSY 209

Qy 61 RIGDTSKDKNRGNLLQICICTGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 210 RIGDTSKDKNRGNLLQICICTGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 259

RESULT 4
US-09-016-366A-12
; Sequence 12, Application US/09016366A
; Patent No. 5955431
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.

```
; APPLICANT: Huang, Chifu
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
; TITLE OF INVENTION: INHIBITORS
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02210-2211
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,366A
; FILING DATE: January 30, 1998
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,090
; FILING DATE: 05-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Plumer, Elizabeth R.
; REGISTRATION NUMBER: 36,637
; REFERENCE/DOCKET NUMBER: B0801/7093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2386 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-016-366A-12

Query Match 100.0%; Score 627; DB 2; Length 2386;
Best Local Similarity 100.0%; Pred. No. 2.4e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGWETWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDRTSY 240

Qy 61 RIGDTSKKDNRGNLLQICITGNRGKWKRCERTHSVQTTSSGSGPFTDVR 110
Db 241 RIGDTSKKDNRGNLLQICITGNRGKWKRCERTHSVQTTSSGSGPFTDVR 290

RESULT 5
US-09-961-403-1
; Sequence 1, Application US/09961403
; Patent No. 6780594
; GENERAL INFORMATION:
; APPLICANT: HE-STUMPP, HOLGER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: KRAETZSCHMAR, JOERN
; APPLICANT: KREFT, BERTHOLT
; APPLICANT: WINTERHAGER, ELKE
; APPLICANT: REGIDOR, PEDRO
; APPLICANT: SCOTTI, SIMONE
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS
; FILE REFERENCE: SCH-1789
; CURRENT APPLICATION NUMBER: US/09/961,403
; CURRENT FILING DATE: 2001-09-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2386

Query Match 100.0%; Score 627; DB 2; Length 2386;
Best Local Similarity 100.0%; Pred. No. 2.4e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGWETWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDRTSY 240

Qy 61 RIGDTSKKDNRGNLLQICITGNRGKWKRCERTHSVQTTSSGSGPFTDVR 110
Db 241 RIGDTSKKDNRGNLLQICITGNRGKWKRCERTHSVQTTSSGSGPFTDVR 290

RESULT 6
US-08-551-356-2
; Sequence 2, Application US/08551356
; Patent No. 5830700
; GENERAL INFORMATION:
; APPLICANT: Irani, Meher
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/551,356
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/349,762
; FILING DATE:
; APPLICATION NUMBER: US/07/998,271
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Holly, Julie A
; REGISTRATION NUMBER: 33-246
; REFERENCE/DOCKET NUMBER: 92-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2446 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-551-356-2

Query Match 100.0%; Score 627; DB 2; Length 2446;
Best Local Similarity 100.0%; Pred. No. 2.5e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGWETWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDRTSY 240

Qy 61 RIGDTSKKDNRGNLLQICITGNRGKWKRCERTHSVQTTSSGSGPFTDVR 110
Db 241 RIGDTSKKDNRGNLLQICITGNRGKWKRCERTHSVQTTSSGSGPFTDVR 290
```

```
;SEQ ID NO:1:
;   LENGTH: 2327
;   5455158-1

Query Match          99.2%; Score 622; DB 6; Length 2327;
Best Local Similarity 99.1%; Pred No. 8.2e-59;
Matches 109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 PIAEKCFDHAAGTSYYVGETWEXPYQGMMVMVDCTCLGEGSGRIITCTSRNRCNDQDTRTSY 60
|       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      153 PIAEKCFDHAAGTPYVVGETWEXPYQGMVMVDCTCLGEGSGRIITCTSRNRCNDQDTRTSY 212
|       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY      61 RIDGTWSKDNKRNGLLOCITGNRGGRGWKCERHTSVQTSSGGSPFTDVR 110
|       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      213 RIDGTWSKDNKRNGLLOCITGNRGGRGWKCERHTSVQTSSGGSPFTDVR 262
|       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
US-08-142-449B-14
; Sequence 14, Application US/08142449B
; Patent No. 5668104
; GENERAL INFORMATION:
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kawano, Genji
; APPLICANT: Sudo, Tetsuo
; APPLICANT: Kojima, Katsuki
; TITLE OF INVENTION: Physiologically Active Protein and
; TITLE OF INVENTION: Hematopoietic Stem Cell Growth Agent
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nels T. Lippert, White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,449B
; FILING DATE: 24-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lippert, Nels T.
; REGISTRATION NUMBER: 25,888
; REFERENCE/DOCKET NUMBER: 1145358-304
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)-819-8582
; TELEFAX: (212) 354-8113
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 188 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-142-449B-14

Query Match          34.0%; Score 213; DB 1; Length 188;
Best Local Similarity 37.6%; Pred No. 1.5e-15;
Matches 35; Conservative 17; Mismatches 39; Indels 2; Gaps 1;

QY      1 PIAEKCFDHAAGTSYYVGETWEXPYQGMMVMVDCTCLGEGSGRIITCTSRNRCNDQDTRTSY 60
|       | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      3 PTDDSCFDPTVSHVAYGDWERMSEGSFKLLCQLGLFGSGHFRCDSSRWCHDNG--VNY 60
|       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
|       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY      61 RIDGTWSKDNKRNGLLOCITGNRGGRGWKCERH 93
|       | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      61 KI GKWRDQGEENGQMMSCTCLGNKGGEFKCDPH 93
|       | | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```


RESULT 10

US-08-836-854-12
; Sequence 12, Application US/08836854
; Patent No. 5824547
; GENERAL INFORMATION:
; APPLICANT: HASHINO, Kimikazu
; APPLICANT: MATSUSHITA, Hideyuki
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: METHOD OF PRODUCTION OF TRANSFECTED CELLS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W. Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,854
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/02425
; FILING DATE: 29-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 317721/1994
; FILING DATE: 29-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: HASHINO=1
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-836-854-12

Query Match 34.0%; Score 213; DB 2; Length 422;
Best Local Similarity 37.6%; Pred. No. 3.8e-15;
Matches 35; Conservative 17; Mismatches 39; Indels 2; Gaps 1;
Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMVMVDTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 286 PTDDSCFDPYTVSHYAVGDEWERMESGFKLLCQCLGFGSGHFRCDSSRWCHDNG--VNY 343
Qy 61 RIGDTWSKDNKGNLLQCTGNGRGWKCERH 93
Db 344 KIGEKWDROGNGOMMSCTCLGNGKGFKCDPH 376

RESULT 11

US-08-836-854-9
; Sequence 9, Application US/08836854
; Patent No. 5824547
; GENERAL INFORMATION:
; APPLICANT: HASHINO, Kimikazu
; APPLICANT: MATSUSHITA, Hideyuki
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: METHOD OF PRODUCTION OF TRANSFECTED CELLS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark

STREET: 419 Seventh Street N.W. Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,854
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/02425
FILING DATE: 29-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 317721/1994
FILING DATE: 29-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: HASHINO=1
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-836-854-9

Query Match 34.0%; Score 213; DB 2; Length 474;
Best Local Similarity 37.6%; Pred. No. 4.4e-15;
Matches 35; Conservative 17; Mismatches 39; Indels 2; Gaps 1;
Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMVMVDTCLGEGSGRITCTSRNRCNDQDRTSY 60
Db 289 PTDDSCFDPYTVSHYAVGDEWERMESGFKLLCQCLGFGSGHFRCDSSRWCHDNG--VNY 346
Qy 61 RIGDTWSKDNKGNLLQCTGNGRGWKCERH 93
Db 347 KIGEKWDROGNGOMMSCTCLGNGKGFKCDPH 379

RESULT 12

US-08-982-597A-20
; Sequence 20, Application US/08982597A
; Patent No. 5932693
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staatz, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 5932693th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,597A
; FILING DATE:

CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,542
FILING DATE: 10-DEC-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-3002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-3117
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-982-597A-20

Query Match 20.3%; Score 127; DB 2; Length 48;
Best Local Similarity 47.4%; Pred. No. 6.9e-07;
Matches 18; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

QY 60 YRIGDTWSKKNRGNLLQICITGNGRGWKCRHTSVQ 97
DB 9 YRIGDQDKQHDGMHMRCTCVGNGRGWTCYAYSQLR 46

RESULT 13
US-09-136-218-20
Sequence 20, Application US/09136218
Patent No. 6083914
GENERAL INFORMATION:
APPLICANT: Santoro, Samuel A.
APPLICANT: Staatz, William D.
TITLE OF INVENTION: Antithrombotic Peptides
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer
STREET: 800 No. 6083914th Lindbergh Blvd.
CITY: St. Louis
STATE: MO
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,218
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/982,597
FILING DATE:
APPLICATION NUMBER: 60/032,542
FILING DATE: 10-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-3002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-3117
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-136-218-20

Query Match 20.3%; Score 127; DB 3; Length 48;

Best Local Similarity 47.4%; Pred. No. 6.9e-07;
Matches 18; Conservative 10; Mismatches 10; Indels 0; Gaps 0;
QY 60 YRIGDTWSKKNRGNLLQICITGNGRGWKCRHTSVQ 97
DB 9 YRIGDQDKQHDGMHMRCTCVGNGRGWTCYAYSQLR 46

RESULT 14
US-08-982-597A-22
Sequence 22, Application US/08982597A
Patent No. 5932693
GENERAL INFORMATION:
APPLICANT: Santoro, Samuel A.
APPLICANT: Staatz, William D.
TITLE OF INVENTION: Antithrombotic Peptides
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer
STREET: 800 No. 5932693th Lindbergh Blvd.
CITY: St. Louis
STATE: MO
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,597A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,542
FILING DATE: 10-DEC-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-3002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-3117
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-982-597A-22

Query Match 17.6%; Score 110.5; DB 2; Length 48;
Best Local Similarity 48.8%; Pred. No. 4.3e-05;
Matches 20; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

QY 51 CNDQDTRTSYRICDTWSKKNRGNLLQICITGNGRGWKCE 91
DB 1 CQDSETGTFTYQIGDSWEKYVH-GVRYQCYCYGRGIGEWHCQ 40

RESULT 15
US-09-136-218-22
Sequence 22, Application US/09136218
Patent No. 6083914
GENERAL INFORMATION:
APPLICANT: Santoro, Samuel A.
APPLICANT: Staatz, William D.
TITLE OF INVENTION: Antithrombotic Peptides
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scott J. Meyer
STREET: 800 No. 6083914th Lindbergh Blvd.
CITY: St. Louis

STATE: MO
COUNTRY: USA
ZIP: 63167
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/136,218
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/982,597
FILING DATE:
APPLICATION NUMBER: 60/032,542
FILING DATE: 10-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meyer, Scott J.
REGISTRATION NUMBER: 25,275
REFERENCE/DOCKET NUMBER: WU-3002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-694-3117
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-136-218-22

Query Match 17.6%; Score 110.5; DB 3; Length 48;
Best Local Similarity 48.8%; Pred. No. 4.3e-05;
Matches 20; Conservative 5; Mismatches 15; Indels 1; Gaps 1;
Qy 51 CNDQDTRTSYRIGDTWSKDNKRNLLQCTGNGRGWKCE 91
Db 1 QDSETGTFTYQIGDSWEKYVH-GVRYQCYCYGRGIGEWHCQ 40

Search completed: November 3, 2004, 23:57:00
Job time : 14.3733 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:54:46 ; Search time 40.3082 Seconds
(without alignments)
884.776 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAKCFDHAAGTSYVVGET.....ERHSTVQTSSGSGPFTDVR 110

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/CTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	627	100.0	259	10	US-09-940-235-4
2	627	100.0	463	14	US-10-144-194A-52
3	627	100.0	642	16	US-10-741-601-354
4	627	100.0	657	16	US-10-741-601-359
5	627	100.0	984	16	US-10-741-601-356
6	627	100.0	2220	15	US-10-236-392-4
7	627	100.0	2296	16	US-10-741-601-363
8	627	100.0	2320	14	US-10-279-733-8
9	627	100.0	2320	15	US-10-236-392-2
10	627	100.0	2328	14	US-10-171-311-64
11	627	100.0	2328	14	US-10-236-031B-70
12	627	100.0	2328	14	US-10-374-979-98
13	627	100.0	2328	15	US-10-182-936A-98

14	627	100.0	2355	14	US-10-144-194A-104
15	627	100.0	2355	15	US-10-360-101-235
16	627	100.0	2355	15	US-10-447-161-3
17	627	100.0	2355	16	US-10-734-564-94
18	627	100.0	2355	16	US-10-741-601-357
19	627	100.0	2355	16	US-10-741-601-366
20	627	100.0	2386	10	US-09-961-403-1
21	627	100.0	2386	16	US-10-741-601-360
22	627	100.0	2386	17	US-10-788-793-206
23	324	51.7	164	9	US-09-925-301-1270
24	244	38.9	231	9	US-09-925-302-548
25	244	38.9	231	10	US-09-925-302-548
26	213	34.0	693	16	US-10-741-601-364
27	213	34.0	793	14	US-10-171-311-62
28	213	34.0	1259	16	US-10-741-601-365
29	213	34.0	1315	16	US-10-741-601-358
30	213	34.0	1341	16	US-10-741-601-355
31	213	34.0	1348	16	US-10-741-601-353
32	201	32.1	343	9	US-09-934-706-1
33	201	32.1	400	9	US-09-934-706-5
34	201	32.1	501	9	US-09-934-706-4
35	165	26.3	163	16	US-10-734-564-115
36	165	26.3	163	17	US-10-770-668-38
37	159.5	25.4	847	16	US-10-741-601-361
38	159.5	25.4	1286	16	US-10-741-601-362
39	156.5	25.0	170	15	US-10-264-049-3220
40	139	22.2	386	15	US-10-344-634-14
41	136.5	21.8	228	15	US-10-344-634-4
42	136.5	21.8	285	15	US-10-344-634-16
43	90.5	14.4	86	15	US-10-424-599-230915
44	75	12.0	1090	15	US-10-149-310-176
45	74.5	11.9	3097	14	US-10-174-677-6

ALIGNMENTS

RESULT 1

US-09-940-235-4
; Sequence 4, Application US/09940235
; Publication No. US20030059921A1
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kammara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION OF SAID
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4

Query Match 100.0%; Score 627; DB 10; Length 259;
Best Local Similarity 100.0%; Pred. No. 1e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGETWEKPYQGMWMDCTCLGEGSGRITCTSRNCNDQDRTSY 60

```
Db 150 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 209
      |||||||
Qy 61 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 110
      |||||||
Db 210 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 259
      |||||||

RESULT 2
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-194A-52

Query Match 100.0%; Score 627; DB 14; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.9e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
      |||||||
Db 217 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 276
      |||||||
Qy 61 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 110
      |||||||
Db 277 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 326
      |||||||

RESULT 3
US-10-741-601-354
; Sequence 354, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match 100.0%; Score 627; DB 16; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.7e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
      |||||||
Db 181 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
      |||||||
Qy 61 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 110
      |||||||
Db 241 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 290
      |||||||

RESULT 4
US-10-741-601-359
```

```
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match 100.0%; Score 627; DB 16; Length 657;
Best Local Similarity 100.0%; Pred. No. 2.8e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
      |||||||
Db 181 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
      |||||||
Qy 61 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 110
      |||||||
Db 241 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 290
      |||||||

RESULT 5
US-10-741-601-356
; Sequence 356, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 356
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-356

Query Match 100.0%; Score 627; DB 16; Length 984;
Best Local Similarity 100.0%; Pred. No. 4.2e-57;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
      |||||||
Db 181 PIAEKFDAAGTSYVVGTEWKPQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
      |||||||
Qy 61 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 110
      |||||||
Db 241 RIGDTSKDNKRNLLQICITGNGRGWKCERHTSVQTTSSGSGPPTDVR 290
      |||||||

RESULT 6
US-10-236-392-4
; Sequence 4, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
```

```
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shalomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladimir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRoche, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shinkets, Richard A
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT APPLICATION NUMBER: US/10/236,392
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130
; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 4
; LENGTH: 2220
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-392-4
Query Match 100.0%; Score 627; DB 15; Length 2220;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMWVDCTCLGSGGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGTEWKPQGMWVDCTCLGSGGRITCTSRNRCNDQDRTSY 240

Qy 61 RIGDTSKDNKGNLLQCICTGNRGKWKERHTSVQTTSSGSGPFTDVR 110
Db 241 RIGDTSKDNKGNLLQCICTGNRGKWKERHTSVQTTSSGSGPFTDVR 290

RESULT 7
US-10-741-601-363
; Sequence 363, Application US/10741601
```

```
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 363
; LENGTH: 2296
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-363
Query Match 100.0%; Score 627; DB 16; Length 2296;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMWVDCTCLGSGGRITCTSRNRCNDQDRTSY 60
Db 181 PIAKCFDHAAGTSYVVGTEWKPQGMWVDCTCLGSGGRITCTSRNRCNDQDRTSY 240

Qy 61 RIGDTSKDNKGNLLQCICTGNRGKWKERHTSVQTTSSGSGPFTDVR 110
Db 241 RIGDTSKDNKGNLLQCICTGNRGKWKERHTSVQTTSSGSGPFTDVR 290

RESULT 8
US-10-279-733-8
; Sequence 8, Application US/10279733
; Publication No. US20030194400A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; APPLICANT: Liu, Cheng
; APPLICANT: Edgington, Thomas, S.
; TITLE OF INVENTION: Targeted Thrombosis
; FILE REFERENCE: 1361.016W01
; CURRENT APPLICATION NUMBER: US/10/279,733
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: US 60/336331
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 2320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-279-733-8
Query Match 100.0%; Score 627; DB 14; Length 2320;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGTEWKPQGMWVDCTCLGSGGRITCTSRNRCNDQDRTSY 60
Db 146 PIAKCFDHAAGTSYVVGTEWKPQGMWVDCTCLGSGGRITCTSRNRCNDQDRTSY 205

Qy 61 RIGDTSKDNKGNLLQCICTGNRGKWKERHTSVQTTSSGSGPFTDVR 110
Db 206 RIGDTSKDNKGNLLQCICTGNRGKWKERHTSVQTTSSGSGPFTDVR 255

RESULT 9
US-10-236-392-2
; Sequence 2, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgees, Catherine, E
; APPLICANT: Casman, Stacie J
```

```

US-10-171-311-64
; Sequence 64, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerssh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-64

```

```

RESULT 11
US-10-236-031B-70
; Sequence 70, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 2328
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-70

```

Query Match	100.0%;	Score 627;	DB 14;	Length 2328;
Best Local Similarity	100.0%;	Pred. No. 1e-56;		
Matches 110;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

Query Match      100.0%; Score 627; DB 15; Length 2320;
Best Local Similarity 100.0%; Pred. No. 1e-56;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
      |||
Db      146 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 205
      |||

QY      61  RIGDTWSKDNKRNGLLQICITGNGRGEWKRCERHTSVQTTSSGSGPFTDVR 110
      |||
Db      206 RIGDTWSKDNKRNGLLQICITGNGRGEWKRCERHTSVQTTSSGSGPFTDVR 255
      |||

RESULT 10

```


US-10-360-101-235

Query Match 100.0%; Score 627; DB 15; Length 2355;
 Best Local Similarity 100.0%; Pred. No. 1.le-56;
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDTCLGEGSGRITCTSRNRCNDQDTRTSY 60
 Db 150 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDTCLGEGSGRITCTSRNRCNDQDTRTSY 209
 Oy 61 RIGDTWSKKDNRGNLLQICITGNRGCEWKCEKERTSVQTTSSGSGPFTDVR 110
 Db 210 RIGDTWSKKDNRGNLLQICITGNRGCEWKCEKERTSVQTTSSGSGPFTDVR 259

Search completed: November 4, 2004, 00:15:28
 Job time : 40.3082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 2004, 23:44:40 ; Search time 161.045 Seconds
(without alignments)
756.808 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAKCFDHAAGTSYVVGET.....ERHSTVQTTSGSGPFTDVR 110

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 6730630 seqs, 1107998698 residues

Total number of hits satisfying chosen parameters: 6730630

Minimum DB seq'length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
- 30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
- 31: /cgn2_6/ptodata/1/paa/US105_COMB.pep.*
- 32: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
- 33: /cgn2_6/ptodata/1/paa/US107_COMB.pep.*
- 34: /cgn2_6/ptodata/1/paa/US108_COMB.pep.*
- 35: /cgn2_6/ptodata/1/paa/US109_COMB.pep.*
- 36: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	627	100.0	259	18	US-09-471-349-4	Sequence 4, Appli
2	627	100.0	259	24	US-09-940-235-4	Sequence 4, Appli
3	627	100.0	259	32	US-10-631-558-4	Sequence 4, Appli
4	627	100.0	379	22	US-09-791-537-53369	Sequence 53369, A
5	627	100.0	463	1	PCT-US02-31287-52	Sequence 52, Appl
6	627	100.0	463	27	US-10-144-194A-52	Sequence 52, Appl
7	627	100.0	463	30	US-10-491-566-52	Sequence 52, Appl
8	627	100.0	642	1	PCT-US03-40977-354	Sequence 354, App
9	627	100.0	642	1	PCT-US03-40978-1066	Sequence 1066, Ap
10	627	100.0	642	33	US-10-741-600-1066	Sequence 1066, Ap
11	627	100.0	642	33	US-10-741-601-354	Sequence 354, App
12	627	100.0	642	33	US-10-796-307-561	Sequence 561, App
13	627	100.0	642	35	US-10-941-087-249	Sequence 249, App
14	627	100.0	642	35	US-10-941-087-257	Sequence 257, App
15	627	100.0	642	36	US-60-568-219-317	Sequence 317, App
16	627	100.0	642	36	US-60-576-801-363	Sequence 363, App
17	627	100.0	642	36	US-60-576-801-373	Sequence 373, App
18	627	100.0	657	1	PCT-US03-40977-359	Sequence 359, App
19	627	100.0	657	1	PCT-US03-40978-1072	Sequence 1072, Ap
20	627	100.0	657	19	US-09-581-651B-36	Sequence 36, Appl
21	627	100.0	657	19	US-09-581-651C-36	Sequence 36, Appl
22	627	100.0	657	33	US-10-741-600-1072	Sequence 1072, Ap
23	627	100.0	657	33	US-10-741-601-359	Sequence 359, App
24	627	100.0	657	33	US-10-796-307-567	Sequence 567, App
25	627	100.0	657	35	US-10-941-087-259	Sequence 259, App
26	627	100.0	657	36	US-60-568-219-323	Sequence 323, App
27	627	100.0	657	36	US-60-576-801-377	Sequence 377, App
28	627	100.0	660	19	US-09-581-651A-1	Sequence 1, Appli
29	627	100.0	675	19	US-09-581-651B-1	Sequence 1, Appli
30	627	100.0	675	19	US-09-581-651C-1	Sequence 1, Appli
31	627	100.0	719	19	US-09-581-651B-37	Sequence 37, Appl
32	627	100.0	719	19	US-09-581-651C-37	Sequence 37, Appl
33	627	100.0	720	19	US-09-581-651B-44	Sequence 44, Appl
34	627	100.0	720	19	US-09-581-651C-44	Sequence 44, Appl
35	627	100.0	727	35	US-10-941-087-250	Sequence 250, App
36	627	100.0	727	36	US-60-576-801-364	Sequence 364, App
37	627	100.0	727	36	US-60-576-801-366	Sequence 366, App
38	627	100.0	843	35	US-10-941-087-251	Sequence 251, App
39	627	100.0	843	36	US-60-576-801-365	Sequence 365, App
40	627	100.0	843	36	US-60-576-801-376	Sequence 376, App
41	627	100.0	843	36	US-60-576-801-378	Sequence 378, App
42	627	100.0	984	1	PCT-US03-40977-356	Sequence 356, App
43	627	100.0	984	1	PCT-US03-40978-1069	Sequence 1069, Ap
44	627	100.0	984	33	US-10-741-600-1069	Sequence 1069, Ap
45	627	100.0	984	33	US-10-741-601-356	Sequence 356, App

ALIGNMENTS

RESULT 1
US-09-471-349-4
; Sequence 4, Application US/09471349
; GENERAL INFORMATION:
; APPLICANT: Sahni, Girish
; APPLICANT: Kumar, Rajesh
; APPLICANT: Roy, Chaiti
; APPLICANT: Rajagopal, Kammar
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE PROTEINS POSSESSING ALTERED
; TITLE OF INVENTION: PLASMINOGEN ACTIVATION CHARACTERISTICS AND A PROCESS FOR THE
; PREPARATION OF SAID PROTEINS
; FILE REFERENCE: 07064/009001
; CURRENT APPLICATION NUMBER: US/09/471,349
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 24

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-471-349-4

Query Match          100.0%; Score 627; DB 18; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 150 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 209

Qy 61 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 210 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 259

RESULT 2
US-09-940-235-4
; Sequence 4, Application US/09940235
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/09/940,235
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-940-235-4

Query Match          100.0%; Score 627; DB 24; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 150 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 209

Qy 61 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 210 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 259

RESULT 3
US-10-631-558-4
; Sequence 4, Application US/10631558
; GENERAL INFORMATION:
; APPLICANT: Kumar, Rajesh
; APPLICANT: Sahni, Girish
; APPLICANT: Roy, Chait
; APPLICANT: Rajagopal, Kamara
; APPLICANT: Nihalani, Deepak
; APPLICANT: Sundaram, Vasudha
```

```
; APPLICANT: Yadav, Mahavir
; TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
; TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
; TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 07064-009002
; CURRENT APPLICATION NUMBER: US/10/631,558
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US/09/940,235
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/471,349
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-558-4

Query Match          100.0%; Score 627; DB 32; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 150 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 209

Qy 61 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 210 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 259

RESULT 4
US-09-791-537-53369
; Sequence 53369, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53369
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53369

Query Match          100.0%; Score 627; DB 22; Length 379;
Best Local Similarity 100.0%; Pred. No. 4.4e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 79 PIAKCFDHAAGTSYVVGWETWPKYQGMWVDDCTCLGEGSGRITCTSRNRCNDQDTRTSY 138

Qy 61 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 110
Db 139 RIGDTSKKNRGNLLQICITGNGRGEWKRCERTHSVQTTSSGSGPPTDVR 188

RESULT 5
PCT-US02-31287-52
; Sequence 52, Application PC/TUS0231287
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
```

; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: PCT/US02/31287
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-31287-52

Query Match 100.0%; Score 627; DB 1; Length 463;
Best Local Similarity 100.0%; Pred. No. 5.5e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 217 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 276
Qy 61 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 277 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 326

RESULT 6
US-10-144-194A-52
; Sequence 52, Application US/10144194A
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-194A-52

Query Match 100.0%; Score 627; DB 27; Length 463;
Best Local Similarity 100.0%; Pred. No. 5.5e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 217 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 276
Qy 61 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 277 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 326

RESULT 7
US-10-491-566-52
; Sequence 52, Application US/10491566
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/491,566
; CURRENT FILING DATE: 2004-04-05
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-491-566-52

Query Match 100.0%; Score 627; DB 30; Length 463;

Best Local Similarity 100.0%; Pred. No. 5.5e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 217 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 276
Qy 61 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 277 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 326

RESULT 8
PCT-US03-40977-354
; Sequence 354, Application PC/TUS0340977
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: PCT/US03/40977
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40977-354

Query Match 100.0%; Score 627; DB 1; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 181 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
Qy 61 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 241 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 290

RESULT 9
PCT-US03-40978-1066
; Sequence 1066, Application PC/TUS0340978
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: PCT/US03/40978
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40978-1066

Query Match 100.0%; Score 627; DB 1; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
Db 181 PIAKCFDHAAGTSYVVGWETWPKYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
Qy 61 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 110
Db 241 RIGDTSKDNKRNLLQCTCTGNRGWKCERHTSVQTTSSGSGPPTDVR 290

```
RESULT 10
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066

Query Match      100.0%; Score 627; DB 33; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
    |||||
Db 181 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
    |||||

QY 61 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
    |||||
Db 241 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290
    |||||

RESULT 11
US-10-741-601-354
; Sequence 354, Application US/10741601
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 354
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-354

Query Match      100.0%; Score 627; DB 33; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
    |||||
Db 181 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
    |||||

QY 61 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
    |||||
Db 241 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290
    |||||

RESULT 12
US-10-796-307-561
; Sequence 561, Application US/10796307
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001509
; CURRENT APPLICATION NUMBER: US/10/796,307
; CURRENT FILING DATE: 2004-03-10
```

```
; NUMBER OF SEQ ID NOS: 44201
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 561
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-796-307-561

Query Match      100.0%; Score 627; DB 33; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
    |||||
Db 181 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
    |||||

QY 61 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
    |||||
Db 241 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290
    |||||

RESULT 13
US-10-941-087-249
; Sequence 249, Application US/10941087
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: Lung Cancer Targets and Uses Thereof
; FILE REFERENCE: CL001546
; CURRENT APPLICATION NUMBER: US/10/941,087
; CURRENT FILING DATE: 2004-09-15
; NUMBER OF SEQ ID NOS: 2100
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-941-087-249

Query Match      100.0%; Score 627; DB 35; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
    |||||
Db 181 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
    |||||

QY 61 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
    |||||
Db 241 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290
    |||||

RESULT 14
US-10-941-087-257
; Sequence 257, Application US/10941087
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: Lung Cancer Targets and Uses Thereof
; FILE REFERENCE: CL001546
; CURRENT APPLICATION NUMBER: US/10/941,087
; CURRENT FILING DATE: 2004-09-15
; NUMBER OF SEQ ID NOS: 2100
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 257
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-941-087-257

Query Match      100.0%; Score 627; DB 35; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 60
    |||||
Db 181 PIAKCFDHAAGTSYVVGGETWKEPKYQGMWVMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
    |||||

QY 61 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 110
    |||||
Db 241 RIGDTSKKDNRGNLLQICITGNGRGWKRCERHTSVQTTSSGSGPPTDVR 290
    |||||
```

```

Db      181  PIAEKCFDHAAGTSYVVGWTKPYQGMWVDCICLGEGRITCTSRNRCNDQDTRTSY 240
Qy      61  RIGDTWSKKDNRGNLLQICICNGRGWKCERHTSVQTTSSGSGPFTDVR 110
Db      241  RIGDTWSKKDNRGNLLQICICNGRGWKCERHTSVQTTSSGSGPFTDVR 290

```

```

RESULT 15
US-60-568-219-317
; Sequence 317, Application US/60568219
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSIVENESS TO STATIN TREATMENT, METHODS OF DETECTION AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001516
; CURRENT APPLICATION NUMBER: US/60/568,219
; CURRENT FILING DATE: 2004-05-06
; NUMBER OF SEQ ID NOS: 28217
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-568-219-317

```

```

Query Match      100.0%; Score 627; DB 36; Length 642;
Best Local Similarity 100.0%; Pred. No. 7.9e-59;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  PIAEKCFDHAAGTSYVVGWTKPYQGMWVDCICLGEGRITCTSRNRCNDQDTRTSY 60
Db      181  PIAEKCFDHAAGTSYVVGWTKPYQGMWVDCICLGEGRITCTSRNRCNDQDTRTSY 240

Qy      61  RIGDTWSKKDNRGNLLQICICNGRGWKCERHTSVQTTSSGSGPFTDVR 110
Db      241  RIGDTWSKKDNRGNLLQICICNGRGWKCERHTSVQTTSSGSGPFTDVR 290

```

Search completed: November 4, 2004, 00:11:22
Job time : 162.045 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 5, 2004, 21:50:51 ; Search time 1549.79 Seconds
(without alignments)
3356.488 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627
Sequence: 1 PIAKCFDHAAGTSYVVGET.....ERHSTVQTTSGSGPFTDVR 110

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+2n.model -DEV=xlp
-Q/cgn2_1/USPTO_gp001_p/US09940235/runat_03112004_174039_11252/app_query.fasta_1.1045
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235@cgn_1_1_5965/runat_03112004_174039_11252 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.px.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	627	100.0	1139	9 HSU42404	U42404 Human fibro
2	627	100.0	2147	6 AX003229	AX003229 Sequence
3	627	100.0	2147	6 BD137021	BD137021 Polypepti
4	627	100.0	2147	9 HSA276395	AJ276395 Homo sapi

5	627	100.0	4321	9 HSM806214	BX538045 Homo sapi
6	627	100.0	7323	11 BV178397	BV178397 sqm10026
7	627	100.0	7502	9 HSM804082	AL832771 Homo sapi
8	627	100.0	7679	6 AR274901	AR274901 Sequence
9	627	100.0	7680	6 AR380744	AR380744 Sequence
10	627	100.0	7680	6 AX277596	AX277596 Sequence
11	627	100.0	7680	6 AX333368	AX333368 Sequence
12	627	100.0	7680	9 HSF1B1	X02761 Human mRNA
13	627	100.0	7803	6 AR051657	AR051657 Sequence
14	627	100.0	7868	9 HSM806653	BX640608 Homo sapi
15	627	100.0	7951	9 HSM806901	BX640802 Homo sapi
16	627	100.0	7951	9 HSM806902	BX64182 Homo sapi
17	627	100.0	8027	6 CQ833991	CQ833991 Sequence
18	627	100.0	8030	9 HSM806170	BX538017 Homo sapi
19	627	100.0	8035	9 HSM806171	BX538018 Homo sapi
20	627	100.0	8044	6 AR454662	AR454662 Sequence
21	627	100.0	8044	6 AX281712	AX281712 Sequence
22	627	100.0	8320	9 HSM806267	BX537590 Homo sapi
23	627	100.0	8411	9 HSM806992	BX640875 Homo sapi
24	627	100.0	8421	9 HSM806903	BX640803 Homo sapi
25	623	99.4	8042	9 HSM806805	BX640731 Homo sapi
26	622	99.2	2192	9 HSA535086	AJ535086 Homo sapi
27	620	98.9	7544	9 HSM803509	AL832202 Homo sapi
28	612	97.6	7705	6 AL4133	AL4133 Fibronectin
29	612	97.6	7705	6 AR034630	AR034630 Sequence
30	612	97.6	7705	6 E01162	E01162 cDNA encodi
31	612	97.6	7705	6 I70110	I70110 Sequence 16
32	612	97.6	7705	6 AR364992	AR364992 Sequence
33	581.5	92.7	3059	10 BC051082	BC051082 Mus muscu
34	580.5	92.6	8329	6 AX402055	AX402055 Sequence
35	580.5	92.6	8329	10 RNFBIRON	X15906 Rat mRNA fo
36	539	86.0	8216	5 XELFERNCT	M77820 Xenopus lae
37	539	86.0	8313	5 BC072841	BC072841 Xenopus l
38	520	82.9	3145	5 AF081127	AF081127 Danio rer
39	520	82.9	8267	5 AF081128	AF081128 Danio rer
40	464.5	74.1	8254	5 AY538257	AY538257 Danio rer
41	461.5	73.6	6816	6 CQ715726	CQ715726 Sequence
C 42	439.5	70.1	171655	2 BX950210	BX950210 Danio rer
C 43	439.5	70.1	193489	5 BX005088	BX005088 Zebrafish
C 44	439.5	70.1	217967	2 BX323049	BX323049 Danio rer
45	412.5	65.8	706	6 CQ731570	CQ731570 Sequence

ALIGNMENTS

RESULT 1	HSU42404	Human fibronectin (FN1)	1139 bp	mRNA	linear	PRI 25-MAR-1999
LOCUS	U42404	Human fibronectin (FN1)	mRNA, N-terminal domains, partial cds.			
DEFINITION	U42404					
ACCESSION	U42404					
VERSION	U42404.1	GI:4096845				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1139)					
AUTHORS	Godfrey,H.P. and Ebrahim,A.A.					
TITLE	Submitted (06-DEC-1995) Henry P. Godfrey, Dept. of Pathology, New					
JOURNAL	York Medical College, Basic Science Building, Valhalla, NY 10595,					
FEATURES	USA					
source	Location/Qualifiers					
	1..1139					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/chromosome="2"					
	/map="2q34"					
	/cell_type="mitogen-activated T lymphocytes"					
gene	1..1139					
CDS	/gene="FN1"					
	<1..>1139					

```
/gene="FN1"
/Note="N-terminal domains of T lymphocyte fibronectin"
/codon_start=2
/product="fibronectin"
/protein_id="AAD00015.1"
/db_xref="GI:4096846"
/translation="GNTYRVGDTYERPKDSMIWDCITCIGAGRGRICTIANRCHGGQ
SYKIGDWRPHETGCMLECVCLNGKGEWTKPIAEKCFDHAAGTSYVVGTEWKP
YQGMVDTCLGEGSGRITCTSRNRCNDQDRTSVRIGDTSKDNKRNGLQICITG
NKGWKECHERTSVQTTSSGSPFDVRAAVQPPHPQPPYCHCVTSGVVYSVGM
QWLKTQGNKMLCTCLNGVSCQPTAVTQTYGNNLNGEBCVLPFTYNGRTFYSCITG
RQDGLHWCSTSNYSQDQKYSFCTDHTLVLTQGGNSGALCHFPFLYNNHNYDCTS
EGRRDNMKWCGTTONYDADQKQFQFCFMAHEBICITNEGVMYRIG"
```

ORIGIN

Alignment Scores:

Pred. No.:	4,056-52	Length:	1139
Score:	627.00	Matches:	110
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-940-235-4_COPY_150_259 (1-110) x HSU42404 (1-1139)

Qy	1	ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	20
Db	236	CCCATAGCTGAGAAGTGTTCATGCTGCTGGACATTCCTATGTGTGGAGAAACG	295
Qy	21	TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGluGlySer	40
Db	296	TGGGAGAAGCCCTACCAAGCTGGATGATGTAGATTGTACTTGCCTGGGAGAGCAGC	355
Qy	41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db	356	GGACGCATCACTTGCACCTTCTAGAAATAGATGATCAACGATCAGGACACAAAGCATCCTAT	415
Qy	61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys	80
Db	416	AGAAATGGAGACACTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	475
Qy	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	476	ACAGCAACGCGGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGAGACCAACATCG	535
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Db	536	AGCGGATCTGGCCCTTCACCGATGTTCTGT	565

RESULT 2

AX003229

LOCUS

DEFINITION

Sequence 2 from Patent WO9931233.

ACCESSION

AX003229

VERSION

AX003229.1

GI:9927082

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Schor,S.L. and Schor,A.M.

AUTHORS

Polypeptides, polynucleotides and uses thereof

TITLE

Patent: WO 9931233-A 2 24-JUN-1999;

JOURNAL

SCHOR SETH LAWRENCE (GB); UNIV DUNDEE (GB)

FEATURES

Location/Qualifiers

1..2147

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

source

ORIGIN

Alignment Scores:

Pred. No.:	7,76e-52	Length:	2147
Score:	627.00	Matches:	110
Percent Similarity:	100.00%	Conservative:	0

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

7.76e-52

627.00

100.00%

100.00%

100.00%

6

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

2147

110

0

0

0

0

US-09-940-235-4_COPY_150_259 (1-110) x AX003229 (1-2147)

Qy	1	ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	20
Db	597	CCCATAGCTGAGAAGTGTTCATGATGCTGCTGGACATTCCTATGTGTGGAGAAACG	656
Qy	21	TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGluGlySer	40
Db	657	TGGGAGAAGCCCTACCAAGCTGGATGATGTAGATTGTACTTGCCTGGGAGAGGACG	716
Qy	41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db	717	GGACGCATCACTTGCACCTTCTAGAAATAGATGATCAACGATCAGGACACAAAGCATCCTAT	776
Qy	61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys	80
Db	777	AGAAATGGAGACACTTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	836
Qy	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	837	ACAGCAACGCGGAGAGAGTGGAGTGTGAGAGGCACACCTCTGTGAGACCAACATCG	896
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Db	897	AGCGGATCTGGCCCTTCACCGATGTTCTGT	926

RESULT 3

BD137021

LOCUS

DEFINITION

Polypeptides, polynucleotides and uses thereof.

ACCESSION

BD137021

VERSION

BD137021.1

GI:23231966

KEYWORDS

JP 2002508179-A/1.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 2147)

Schor,S.L. and Schor,A.M.

AUTHORS

Polypeptides, polynucleotides and uses thereof

TITLE

Patent: JP 2002508179-A 1 19-MAR-2002;

JOURNAL

UNIVERSITY OF DUNDEE

COMMENT

OS

Homo sapiens (human)

PN

JP 2002508179-A/1

PD

19-MAR-2002

PF

15-DEC-1998

JP 2000539133

PR

16-DEC-1997

GB 9726539.1

PI

SETH LAWRENCE SCHOR,ANA MARIA SCHOR

PC

C12N15/09,A61K38/00,A61P17/02,C07K14/78,C07K16/18,C12N5/10,PC

C12P21/06,

PC

C12Q1/68,G01N33/574,C12N15/00,A61K37/02,C12N5/00

CC

Polypeptides, polynucleotides and uses thereof. FH

Key

Location/Qualifiers

FT

source

1..2147

/organism="Homo sapiens (human)".

FEATURES

source

1..2147

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.:	7,76e-52	Length:	2147
Score:	627.00	Matches:	110
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-09-940-235-4 COPY 150 259 (1-110) X BD137021 (1-2147)

Qy	1	ProileAlaGluLysCysaphAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	20
Db	597	CCCATAGCTGAGAAAGTGTTTTGATCATGCTGTGGACTTCTATGTGTCGAGAAACG	656
Qy	21	TrpGluLysProTyrGlnGlyTTrpMetMetValAspCysThrCysLeuGlyGluGlySer	40
Db	657	TGGAGAGAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCTGGGAGAAAGCACG	716
Qy	41	GlyArgileThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db	717	GGACGCATCATCTTGACATTCAGAAATAGATGCAAGATCAGACACAAAGSACATCCAT	776
Qy	61	ArgileGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Db	777	AGANTTGGAGACACCTGGAGACAAAGGNTANTCGAGGAAAACCTGCTCAGTGCATCTGC	836
Qy	81	ThrGlyAsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	837	ACAGSACACGCGAGGAGAGTGGAAAGTGTGAGAGSACACACCTCTGTGAGACCACATCG	896
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Db	897	AGCGGATCTGGCCCTTCACCGATGTTCT	926

RESULT 4						PRI 06-JAN-2001
HSAZ76395						
LOCUS	HSAZ76395			2147 bp	mRNA	linear
DEFINITION	Homo sapiens mRNA for MSF-FN70 (FN gene).					
ACCESSION	AJ276395					
VERSION	AJ276395.1			GI:12053816		
KEYWORDS	FN gene; migration stimulating factor FN70.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					

REFERENCE	1	Schor, S.L., Schor, A.M., Seneviratne, K., Kay, R., Ellis, I., Baillie, R. and Clausen, J.
AUTHORS		
TITLE		Migration stimulating factor (MSF): A novel transcription variant of the fibronectin gene
JOURNAL		Unpublished
REFERENCE	2 (bases 1 to 2147)	Schor, S.L.
AUTHORS		
TITLE		Direct Submission
JOURNAL		Submitted (06-MAR-2000) Schor S.L., The Dental School, University of Dundee, Park Place, Dundee DD1 4HR, Scotland
FEATURES		Location/Qualifiers
SOURCE		1. .2147

```

gene
1..2147
/genes="FN"
57..1985
/genes="FN"
/notes="70 kDa truncated isoform of human fibronectin"
/codon_start=1
/product="migration stimulation factor FN70"
/protein_id="CAC20427.1"
/db_xref="GI:12053817"
/db_xref="GOA:P02751"
/db_xref="Swiss-Prot:P02751"
/translation="MLRGPGGLLLAVQCLGTAVPSTGASKSKRAQAAQMVQPQSPVA
VSQSKPCGYDNKHYQINQWERTYLGNALVCTCYGSGRGFNCSKPEASEATCFDKYT
GNTYRGDTYERPKDSMWDCICGAGRVCISCTIANRCHGGQSYKIGDTWRRPHEP
GGYMEICVLGNGDWTCPKIAEKPDHAAGTSYVVGTEWKPYQGMWVDCITLGE
GSGRIICTTSNRGNQDQTSYRIGDTPWASKDNKGNLLLOCITGNGRGWEKCKRHTSV

```

QTTSSGSGPPTDRAAYVQPQHPQPPYCHGCVTDSGVVSYVGKWLKCTQGNKQMLCT
CLNGVSQBTAVTQYGGNSNGPCVLPYTYTYNDRTSTTSNTEQKQKSYCFDHTVVL
VQTRGNSGALCHFPFLYNNHNYTCTSEGRNDNMKCCTQNYQADQKGFPCPMWAA
HEBICTTNGEYVRI GDQDKHDHGMHMECTCVGNGRGWTCIATYSQLRQDCIVDI
TYANNVTFRKREGEHMLNCTCFGQGRKWKCDPVQCDQSTETFYQIGDSNEKYVH
GVRYQCVYCGRGI GEWHQCPQLQTYPSSSGPVEVFI TETPSQNSHP IQWNAPOFPHS
KYTLRWBPVSYI PPPNT.GY²¹

ORIGIN

Alignment Scores:		
Pred. No.:	7,76e-52	Length: 2147
Score:	627.00	Matches: 110
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0

US-09-940-235-4 COPY 150 259 (1-110) x HSA276395 (1-2147)

[illegible]

Qy	61	ArgIleGlyAspThrTrpSerLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Db	777	AGAATTGGGACACCTGGAGCAAGAAGTAATTCGAGGAAACCTGCTCCAGTGCATCTGC	836
Qy	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	837	ACAGGCACCGCCCGAGAGAGTGAAGAGTGAAGAGCACACCTCTGTGCAGACCACATCG	896
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Db	897	AGCGGATCTGGCCCTTCCACCAATGTCGT	926

RESULT 5	
LOCUS	HSM806214
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686B18150 (from clone DKFZp686B18150).
ACCESSION	BX538045
VERSION	BY538045.1
	GI:31874156
	4321 bp mRNA linear PRI 17-JUN-2003

KEYWORD	SOURCE	ORGANISM
.	Homo sapiens (human)	
	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	

REFERENCE	1. (BASES 1 TO 4321)
AUTHORS	Bloecher, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
TITLE	Direct Submission
JOURNAL	Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GBF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKZPp686B18150) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de. Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

```

FEATURES
source
Location/Qualifiers
1. .4321
/organism="Homo sapiens"

```

```
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/map="2q35"  
/clone="DKFZp686B18150"  
/tissue_type="human colon endothel primary cell culture"  
/clone_lib="686 (synonym: hicc3). Vector psporti_sfi; host  
DH10B; sites SfiIA + SfiIB"  
/dev_stage="adult"  
1. 4321  
/gene="DKFZp686B18150"  
31. 2280  
/gene="DKFZp686B18150"  
/note="fibronectin precursor, C-terminus truncated,  
differentially spliced"  
/codon_start=1  
/product="hypothetical protein"  
/protein_id="CAD97984.1"  
/db_xref="GI:31874157"  
/translation="MAGGPRRLCTGGEGTPGAAGKRGPAATTSLVLCIPSVPPVP  
PPTLWPPSRORPPGIRRRDLRRRLREANLVATCLPVRASLPHRLNMLRGPGLL  
LLAVLCGTAVSTGASKRKQAQOMVQPSVAVSQKPGCYDNGKHQINQWERT  
YAGNALVCTCGSGFNCESPEAEETCFDKYGTNTYRGDTYERPKDSMIWDCTCI  
CAGGRISCTTANRHEGQSYKIGDTWRPHETGGYMLECVLNGKEWTKCPIAE  
KCFHAAGTSYVGTWPKPYQMMWVDCITLGEGRITCTSRNRCNDQDRTSYRI  
GDTWSKDNRLNLCICTNGRGEWKCEHTSVOTSSGSPFTDRAAVYQPOPH  
PQPPYCHVTDSGVYSVMOWLKTQGNKQLCTCLNGVSCQETAVTQYGNSEE  
PCVLPFTYNGRTFYCTEGHQDGLHWCSTTSNYQDQKYSFCTDHTVLVQRGNSN  
GALCHFPFLYNNHTDCTSEGRDNMKWGTQNYDADQKFGFCPMAHEBICITNE  
GVMYRIGDQWKDHDMHMRCTCVNGRGEWTCIAYSQLRQDCITVDITVNVNFTF  
KRHEBGLMNCCTCFQGRGWRKCDPVQCDQDSETGTFTYIGDSWKEVHGVRVQCYCY  
GRGIGEWHCPLQTYTPSSSGPVEVFTETPSPNSHPIQWNAQPSHISKYILRWRPV  
SIPRNLGY"  
4297  
/polyA_site  
/gene="DKFZp686B18150"  
  
Alignment Scores:  
Pred. No.: 1,59e-51 Length: 4321  
Score: 627.00 Matches: 110  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
  
US-09-940-235-4_COPY_150_259 (1-110) x HSM806214 (1-4321)  
  
Qy 1 ProlleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20  
Db 847 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGGTGGAGAAACG 906  
  
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40  
Db 907 TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAACG 966  
  
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60  
Db 967 GGACGATCCTTGTGACCTTCTAGAAATAGATGCAACGATCAGACACAGGACATCCTAT 1026  
  
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80  
Db 1027 AGAATTGGAGACACCTGGGACCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1086  
  
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100  
Db 1087 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGCACACCTCTGTGCAGACCATCG 1146  
  
Qy 101 SerGlySerGlyProPheThrAspValArg 110  
Db 1147 AGCGGATCTGGCCCTTACCCGATGTTTGGT 1176  
  
RESULT 6  
BV178397 HSM804082 7323 bp DNA linear STS 10-JUN-2004  
LOCUS BV178397 Homo sapiens mRNA; cDNA DKFZp686B18150 (from clone DKFZp686B18150).
```

```
DEFINITION sqm100268 Human DNA (Sequenom) Homo sapiens STS genomic, sequence  
tagged site.  
ACCESSION BV178397  
VERSION BV178397.1 GI:48014736  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 7323)  
AUTHORS Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,  
Cantor,C.R. and Braun,A.  
TITLE Large-Scale Validation of Single Nucleotide Polymorphisms in Gene  
Regions  
JOURNAL Genome Res. (2004) In press  
COMMENT Contact: Andreas Braun  
Pharmaceuticals division  
Sequenom, Inc.  
3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 7323  
FEATURES  
Location/Qualifiers  
source  
1..7323  
/organism="Homo sapiens"  
/mol_type="genomic DNA"  
/db_xref="taxon:9606"  
/clone_lib="Human DNA (Sequenom)"  
STS  
1..7323  
ORIGIN  
  
Alignment Scores:  
Pred. No.: 2,74e-51 Length: 7323  
Score: 627.00 Matches: 110  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0  
  
US-09-940-235-4_COPY_150_259 (1-110) x BV178397 (1-7323)  
  
Qy 1 ProlleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20  
Db 52 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGGTGGAGAAACG 111  
  
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40  
Db 112 TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAACG 171  
  
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60  
Db 172 GGACGATCCTTGTGACCTTCTAGAAATAGATGCAACGATCAGACACAGGACATCCTAT 231  
  
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80  
Db 232 AGAATTGGAGACACCTGGGACCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 291  
  
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100  
Db 292 ACAGGCAACGGCCGAGGAGAGTGAAGTGTGAGAGCACACCTCTGTGCAGACCATCG 351  
  
Qy 101 SerGlySerGlyProPheThrAspValArg 110  
Db 352 AGCGGATCTGGCCCTTACCCGATGTTTGGT 381  
  
RESULT 7  
HSM804082 7502 bp mRNA linear PRI 13-MAY-2003  
LOCUS HSM804082 Homo sapiens mRNA; cDNA DKFZp686B197 (from clone DKFZp686B197).
```

ACCESSION AL832771
VERSION AL832771.1 GI:21733353
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 7502)
Ansong, W., Krieger, S., Regiert, T., Rittmüller, C., Schwager, B., Wewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr. 1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by ENBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp686B197) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.
FEATURES
source
1..7502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686B197"
/tissue type="human cervix"
/clone lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiIA + SfiIB"
/dev_status="adult"
polyA signal 7469..7474
polyA_site 7488
ORIGIN
Alignment Scores:
Pred. No.: 2,81e-51 Length: 7502
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-940-235-4_COPY_150_259 (1-110) x HSM804082 (1-7502)
QY 1 ProilleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 808 CCCATAGCTGAGAGGTTTGTATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 867
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
DB 868 TGGGAGAAGCCCTACCAAGCTGGATGGTAGATTGTCTTCCCTGGGAGAGGCAGC 927
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 928 GGACGCATCATCTTGCACCTCTTAGAAATAGATGCAACGATCAGGACACCAAGCATCTAT 987
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 988 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1047
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 1048 ACAGCAACGGCCGAGAGAGTGAAGTGTGAGAGCACACCTCTGTGCAGACCATCG 1107
QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 1108 AGCGGATCCGGCCCTTACCAGATGTTTCGT 1137
RESULT 8

AR274901
LOCUS AR274901 7679 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 38 from patent US 6506607.
ACCESSION AR274901
VERSION AR274901.1 GI:29707451
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7679)
AUTHORS Shyjan, A.W.
TITLE Methods and compositions for the identification and assessment of prostate cancer therapies and the diagnosis of prostate cancer
JOURNAL Patent: US 6506607-A 38 14-JAN-2003;
FEATURES
source 1..7679
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,88e-51 Length: 7679
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-940-235-4_COPY_150_259 (1-110) x AR274901 (1-7679)
QY 1 ProilleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 461 CCCATAGCTGAGAGGTTTGTATCATGCTGCTGGGACTTCCTATGTGGTCGGAGAAACG 520
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
DB 521 TGGGAGAAGCCCTACCAAGCTGGATGGTAGATTGTCTTCCCTGGGAGAGGCAGC 580
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 581 GGACGCATCATCTTGCACCTCTTAGAAATAGATGCAACGATCAGGACACCAAGCATCTAT 640
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 700
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 701 ACAGCAACGGCCGAGAGAGTGAAGTGTGAGAGCACACCTCTGTGCAGACCATCG 760
QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 761 AGCGGATCTGGCCCTTACCAGATGTTTCGT 790
RESULT 9
AR380744
LOCUS AR380744 7680 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1289 from patent US 6607879.
ACCESSION AR380744
VERSION AR380744.1 GI:40088378
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 7680)
AUTHORS Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.
TITLE Compositions for the detection of blood cell and immunological response gene expression
JOURNAL Patent: US 6607879-A 1289 19-AUG-2003;
FEATURES
source 1..7680
/organism="unknown"
/mol_type="genomic DNA"

Db 701 ACAGCAACGCCGAGGAGAGTGGAGTGTGAGAGCACACCTCTCTGAGACACCATCG 760

Qy 101 SerGlySerGlyProPheThrAspValArg 110
 |||||
 Db 761 AGCGGATCTGGCCCTTCACCGATGTCGT 790

RESULT 12

HSF1B1 7680 bp mRNA linear PRI 01-OCT-1999

LOCUS Human mRNA for fibronectin (FN precursor).
 ACCESSION X02761 K00055 K00799 K02273 X00307 X00739
 VERSION X02761.1 GI:31396
 KEYWORDS alternate splicing; fibronectin.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHOR Kornblihtt, A.R., Vibe-Pedersen, K. and Baralle, F.E.
 TITLE Isolation and characterization of cDNA clones for human and bovine fibronectins
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 80 (11), 3218-3222 (1983)
 MEDLINE 83221567
 PUBMED 6304699

REFERENCE 2
 AUTHOR Kornblihtt, A.R., Vibe-Pedersen, K. and Baralle, F.E.
 TITLE Human fibronectin: molecular cloning and evidence for two mRNA species differing by an internal segment coding for a structural domain
 JOURNAL EMBO J. 3 (1), 221-226 (1984)
 MEDLINE 84158533
 PUBMED 6200322

REFERENCE 3
 AUTHOR Kornblihtt, A.R., Vibe-Pedersen, K. and Baralle, F.E.
 TITLE Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats
 JOURNAL Nucleic Acids Res. 12 (14), 5853-5868 (1984)
 MEDLINE 84272258
 PUBMED 6462919

REFERENCE 4
 AUTHOR Kornblihtt, A.R., Umezawa, K., Vibe-Pedersen, K. and Baralle, F.E.
 TITLE Primary structure of human fibronectin: differential splicing may generate at least 10 polypeptides from a single gene
 JOURNAL EMBO J. 4 (7), 1755-1759 (1985)
 MEDLINE 85284965
 PUBMED 2992939

COMMENT On or before Jan 13, 2004 this sequence version replaced gi:182680, gi:182681, gi:182698.

FEATURES
 source
 1..7680
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 <1..6988
 /codon_start=2
 /product="fibronectin precursor"
 /protein_id="CAA26536.1"
 /db_xref="GI:31397"
 /db_xref="GOA:P02751"
 /db_xref="Swiss-Prot:P02751"
 /translation="KSRQAQMQVQSPVAVSQSGKGYDNGKHVQINQWERTVLG
 NVLVCYSGSRGNCSEKPEAEETCFDKYTGNTYRVGDTYERPKDSIMWDCICGAG
 RRAISCTIANRCHGGQYKIGDWRRHRETGYMLECVCLNGKGEWCKTIAEKC
 DHAAGTSVVGTEWEKPYQGMWMDVCTCLGEGSGRITCTSRNCRNDQDTRTSYRGDT
 WSKDNRLNLQCICTGNRGWCKERTSVQTSSSGSPFDVRAAVTQYQPHQPP
 PYGHCVDGTVYVSGMQLTKQNKQLMCTCLNGVSCQETAVTQYVGNLNGPCV
 LPFYNGRTFYSCTEGGQDHLWCSTTSNVEDOKYSFCTDHTVLVOTQGSNGAL
 YRFFPLYNHNYTCTSGRRDNMKWCGTQNTYQADQKRFQCFMAAHEICTINGVM
 YRIGQMDQNDHMGKCTCVNGRGKCTIAYSQLRDQCIVDDITNVNDTHFRH
 IEGHMLNCTQGGQRGKWDVQDQCSQSETGYIGDSWEKYGVRQYCYGRG
 IEGWHQCLQTPYSGSGVEVFTITPQNSHPIQWNAPOQSHISKTLRWPKNSV
 GRWKEATIPGLHNSYTIKGLKPGVYEGQLISIQYGHQEVTRFDFTTSTPTVSN
 TVTGTPTPPSLVATSESVTEITASSFVSVWSADTVSGFRFVELSEBEGDEPQYLD

CDS

LPSTATSNIPDLLPKRKIVNVYQISEDEQSLILSTSTQTPADAPPDPDTVDQDDT
 SVWRSRQPAITGYRIYVSPSEGSSELTNLPETANSVTLSDLQFGVQYNTIYAV
 ENEQSTPVVIOQETTPRSDTPSPRDLQFVEDVDVKVTIMWTPEASVTGYRVDV
 IPVNLPEHGRLPISRNTFAEVTGSPGVYFKVFAVSHGRESKPLTAQQTTLKDA
 PTNLOFNETDSTLVLRWTPPRAQITGYRLTGLTRGGOPROYNVGSVSKYPLNLQ
 PASBYTVSLVAKGNOESPATGVTTLQPGSSIPPYNTEVTETIIVTWPARGF
 KLGRPSQGEAPREVTSDSGSIIVSGLTPGVBYVTIQLRDGQSDAPVANKVTP
 LSPPTNLHLEAPDTCGLTVSMERSTTPTDITGYRTITPTNGQGSLSLEWHADQSS
 CFPDNLSPGLVNSVYTKDKKESVPSDIIIPAVPPTDLFTNIGPTMTWTPAP
 PSIDLTNLFVNSPVKNEEDVAELSISSDNALVLTNLLPGTEYVYVSVSSVEOES
 TPLRGOKTGLDPTGIDFSDITANSFTVHIAPIRATITGYRIHHEHESGSPREDR
 VPHRNSIHLTPTGTEYVSVIVALNGHESPELLICQQTSDVDFRDLVRAATPTS
 LLISDAPAVTVRYRTIYGETGNSPVQESTFGSKSTATISGLKPGVDYITTVAV
 TCRGDSPASSKPTISINRTEIDKPSQMVTDQDINSISVKWLPSSPVTGYRTTPK
 NPGPPTKTAGPDQTEMTEIGLQPTVEYVYVYAQNPSESQPLVQTAVTNIDRPG
 LAFTDVDSIKIAWSPQGVSRVYTSYPEDGIELFPAPDGEDTAELQGLARG
 SEYVSVALHDDMESQPLIGTQSTAIAPATDLKFTQVTPYLSAQWTPNVQLTGYR
 VRVTPREKTGPMKEINLAPDSSSVVSGLMVATKEYSVVALKDTLSRAQGVTTL
 ENVSPRRARVTDATETITISWRTKTETITGQVDAVPAANGQTPITORTIKPVSYT
 ITGLQPTDYKIYLTNDNARSPPVIDASTAIDAPSNLRFLATPNSILLWQSPR
 ARTGIIKYEKPGSPREVPRPGVTEATITGLPEGYTYIVYIALKNKQSKPL
 IGRKTDLPOLVTLPHNHLGPEILDVPTVOKTPEVTHPGVDTGNGIOLPCTSGQ
 PSVQOMIEFEHGFRTTPTATPIRHRPRYPYPNVQGEALSQTTISWAPFQDTSEY
 IISCHPVGTDEEPLQFRVPTSATILGLTRGATYNIIVALKDQQRHKRVEVTV
 GNSVNEGLNQTDSCFPTVSHYAVGDEMERMSSESKLCCQLGFGSGHFRCDSS
 QWKEYLGAIGKICCTCFGQGRWRCNCRPRGSPSPBEGTTGQSNQVYSQRYHQTWN
 VNCPTFCFPLDVQADREDSRE"

14..6985
 /product="fibronectin precursor"
 /note="region of internal homology I (5 subunits); domain with fibrin-, heparin-, s.aureus-binding activity"
 842..952
 /note="region of internal homology I (1 subunit); domain with collagen-binding activity"
 953..1327
 /note="region of internal homology II (2 subunits) domain with collagen-binding activity"
 1328..1744
 /note="region of internal homology I (3 subunits) domain with collagen-binding activity"
 1745..2020
 /note="region of internal homology III (1 subunit)"
 2075..5893
 /note="region of internal homology III (14 subunits)"
 2347..3175
 /note="domain with DNA-binding activity"
 4490..4501
 /note="cell binding site"
 4811..5080
 /note="ED region for alternative splicing"
 5081..5893
 /note="domain with heparin-binding activity"
 5894..6160
 /note="III CS region for alternative splicing"
 6161..6373
 /note="region of internal homology III (1 subunit)"
 6442..6838
 /note="region of internal homology I (3 subunits); domain with fibrin-binding activity"
 7679

polya_site

ORIGIN

Alignment Scores:

Pred. No.: 2,88e-51
 Score: 627.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 9

Length: 7680
 Matches: 110
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

Db	810	CCCATAGCTGAGAACTGTTTTCATCATGCTGCTGGGACTTCCTATGCTGTCGAGAAACG	869
Qy	21	TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer	40
Db	870	TGGGAGAAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCCTGGGAGAGGCACG	929
Qy	41	GlyArgileThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db	930	GGACGCATCATTTGCACCTTCTAGAAATAGATGCACGATCAGGACACAAAGGACATCCTAT	989
Qy	61	ArgileGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysileCys	80
Db	990	AGAAATTGGAGACACCTGGAGCAAGAGGATATTCGAGGAAACCTGCTCCAGTGCATCTGC	1049
Qy	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	1050	ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG	1109
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Db	1110	AGCGATCTGGCCCCCTTACCGATGTTTGT	1139

Search completed: November 6, 2004, 02:55:21
Job time : 1559.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 5, 2004, 21:47:11 ; Search time 176.113 Seconds
(without alignments)
3278.783 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAEKCFCFHAGTGVVGET.....ERHSTVQTTSGSGPFTDVR 110

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_03112004_174038_11244/app_query.fasta_1.1045
-DB=N Geneseq_23Sep04 -OPT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09940235 -CGN_1_1063@runat_03112004_174038_11244 -NCFU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_23Sep04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	627	100.0	777	3	AAA37632 Human fib
2	627	100.0	1541	3	AAA37644 Chimeric
3	627	100.0	2096	3	AAA37643 Chimeric
4	627	100.0	2127	10	ADD18477 Human pro
5	627	100.0	2147	2	AAX81299 Human mig
6	627	100.0	3540	12	ADP75954 Human min

7	627	100.0	4080	12	ADP75958	Adp75958 Human leu
8	627	100.0	4295	8	ACC72037	Acc72037 BC007770 g
9	627	100.0	4860	3	AAA35009	Aaa35009 Human ade
10	627	100.0	6988	9	ACD06169	Acd06169 Human cdn
11	627	100.0	7049	12	ADN95947	Adn95947 Human NOV
12	627	100.0	7361	9	ACD06170	Acd06170 Human cdn
13	627	100.0	7550	8	ACC00412	Acc00412 Human cel
14	627	100.0	7679	10	ADB31322	Adb31322 Testoster
15	627	100.0	7680	2	AAT17551	Aat17551 Human low
16	627	100.0	7680	3	AAF21131	Aaf21131 Human fib
17	627	100.0	7680	5	ABA82689	Aba82689 Fibronect
18	627	100.0	7680	6	ABL67540	AbL67540 Thyroid c
19	627	100.0	7680	6	ABT11082	Abt11082 Human bre
20	627	100.0	7680	8	ACF03878	Acf03878 Human fib
21	627	100.0	7680	8	ABX10391	Abx10391 DNA encod
22	627	100.0	7680	8	ACC46009	Acc46009 Human fib
23	627	100.0	7680	8	ACF12859	Acf12859 Human cer
24	627	100.0	7680	8	ACA64817	Aca64817 Human fib
25	627	100.0	7680	8	ACA64819	Aca64819 Human fib
26	627	100.0	7680	9	ADB70377	Adb70377 Fibronect
27	627	100.0	7680	10	ADB98703	Adb98703 Human fib
28	627	100.0	7680	10	ADD18771	Add18771 Human dis
29	627	100.0	7680	10	AD82499	Ad82499 Human DNA
30	627	100.0	7680	10	ABZ96825	Abz96825 Human nuc
31	627	100.0	7680	11	ADI31963	Adi31963 Human cdn
32	627	100.0	7680	11	ABD20674	Abd20674 Human pul
33	627	100.0	7680	12	ADJ37156	Adj37156 Human mal
34	627	100.0	7680	12	ADP13486	Adp13486 Renal cel
35	627	100.0	7795	10	ADJ56196	Adj56196 Zebrafish
36	627	100.0	7803	2	AAQ70009	Aaq70009 Fibrinoge
37	627	100.0	7867	4	AAI57803	Aai57803 Human pol
38	627	100.0	8027	11	ADP64998	Adp64998 Human fib
39	627	100.0	8027	12	ADG89565	Adg89565 Human fib
40	627	100.0	8039	3	AAC89889	Aac89889 Human FN
41	627	100.0	8044	6	AAS94866	Aas94866 Human DNA
42	627	100.0	8044	12	ADI61767	Adi61767 Human cdn
43	627	100.0	8062	4	AAI57802	Aai57802 Human pol
C 44	627	100.0	8066	4	AAI59591	Aai59591 Human pol
C 45	627	100.0	8066	4	AAI59589	Aai59589 Human pol

ALIGNMENTS

RESULT 1

AAA37632
ID AAA37632 standard; DNA; 777 BP.

XX AC AAA37632;

XX DT 13-OCT-2000 (first entry)

XX DE Human fibronectin coding sequence fragment.

XX KW Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
XX KW plasminogen; human; fibronectin; thrombolytic therapy;
XX KW cardiovascular disorder; fibronectin; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1..777

FT FT /*tag= a

FT FT /product= "fibronectin"

FT FT /partial

XX PN EP1024192-A2.

XX XX

XX PD 02-AUG-2000.

XX PF 23-DEC-1999; 99EP-00310541.

XX PR 24-DEC-1998; 98IN-DE003825.

```
PA (COUL ) CSIR COUNCIL SCI IND RES.
XX
XX
PI Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
PI Yadav M;
XX
XX WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX domains of human fibronectin.
XX
XX Example 3; Fig 6; 58pp; English.
XX
XX This sequence represents a human fibronectin coding sequence fragment,
XX containing fibrin binding domains. The invention relates to a hybrid
XX plasminogen activator (PA) comprises a polypeptide fusion between
XX streptokinase (SK), which are capable of plasminogen (PG) activation, and
XX fibrin binding regions of human fibronectin, which are from fibrin
XX binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
XX ability to bind with fibrin independently and also characteristically
XX retains a PG activation ability which becomes evident only after a
XX pronounced duration, or lag, after exposure of the PA to a suitable
XX animal or human PG. The hybrid streptokinase-fibrin binding domain
XX polypeptides are useful in thrombolytic therapy for various kinds of
XX cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
XX well as kinetics of plasminogen activation that are distinct from that of
XX natural streptokinase in being characterised by a temporary delay, or lag
XX of several minutes in the natural rate of the catalytic conversion of
XX plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
XX can bind tightly with fibrin in blood clots soon after introduction into
XX the vascular system without significantly activating the circulating
XX blood plasminogen to plasmin, thus aiding in the localisation of the
XX plasminogen activation process to the site of pathological thrombus. This
XX overcomes systemic plasminogen activation encountered during clinical use
XX of streptokinase
XX
XX Sequence 777 BP; 214 A; 166 C; 230 G; 167 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 2.63e-58 Length: 777
XX Score: 627.00 Matches: 110
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 3 Gaps: 0
XX
XX US-09-940-235-4_COPY_150_259 (1-110) x AAA37644 (1-777)
XX
XX Qy 1 ProfileAlaGluLysCysPheAspHisAlaIaGlyThrSerTyrValValGlyGluThr 20
XX Db 448 CCCATAGCTGAGAAAGTGTGTTGATCATGCTGTGGGACTTCTATGTGGTCGGAGAAACG 507
XX
XX Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
XX Db 508 TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGACTTGGCTGGGAGAGGCAGC 567
XX
XX Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
XX Db 568 GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 627
XX
XX Qy 61 ArgIleGlyAspThrTyrSerIlyslsAspAsnArgGlyAsnLeuGlnCysIleCys 80
XX Db 628 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 687
XX
XX Qy 81 ThrGlyAsnGlyAtrGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
XX Db 688 ACAGGCACCGCCCGGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 747
XX
XX Qy 101 SerGlySerGlyProPheThrAspValArg 110
XX Db 748 AGCGGATCTGCCCCCTTCACCGATGTTCTG 777
XX
XX RESULT 2
XX
XX AAA37644
XX ID AAA37644 standard; DNA; 1541 BP.
XX AC AAA37644;
XX
XX 15-SEP-2003 (revised)
XX 13-OCT-2000 (first entry)
XX
XX Chimeric SK-FBD coding sequence.
XX
XX Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;
XX plasminogen; human; fibronectin; thrombolytic therapy;
XX cardiovascular disorder; ss.
XX
XX Streptococcus dysgalactiae subsp. equisimilis.
XX Homo sapiens.
XX Chimeric.
XX
XX EP1024192-A2.
XX
XX 02-AUG-2000.
XX
XX 23-DEC-1999; 99EP-00310541.
XX
XX 24-DEC-1998; 98IN-DE003825.
XX
XX (COUL ) CSIR COUNCIL SCI IND RES.
XX
XX Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;
XX Yadav M;
XX
XX WPI; 2000-516032/47.
XX
XX Hybrid streptokinase-fibrin binding domain polypeptides useful for
XX thrombolytic therapy comprises a streptokinase fused with fibrin binding
XX domains of human fibronectin.
XX
XX Disclosure; Fig 17b; 58pp; English.
XX
XX This sequence represents a chimeric streptokinase-fibrin binding domain
XX (SK-FBD) protein coding sequence. The invention relates to a hybrid
XX plasminogen activator (PA) comprises a polypeptide fusion between
XX streptokinase (SK), which are capable of plasminogen (PG) activation, and
XX fibrin binding regions of human fibronectin, which are from fibrin
XX binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the
XX ability to bind with fibrin independently and also characteristically
XX retains a PG activation ability which becomes evident only after a
XX pronounced duration, or lag, after exposure of the PA to a suitable
XX animal or human PG. The hybrid streptokinase-fibrin binding domain
XX polypeptides are useful in thrombolytic therapy for various kinds of
XX cardiovascular disorders. The hybrids have enhanced fibrin selectivity as
XX well as kinetics of plasminogen activation that are distinct from that of
XX natural streptokinase in being characterised by a temporary delay, or lag
XX of several minutes in the natural rate of the catalytic conversion of
XX plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins
XX can bind tightly with fibrin in blood clots soon after introduction into
XX the vascular system without significantly activating the circulating
XX blood plasminogen to plasmin, thus aiding in the localisation of the
XX plasminogen activation process to the site of pathological thrombus. This
XX overcomes systemic plasminogen activation encountered during clinical use
XX of streptokinase
XX
XX Sequence 1541 BP; 497 A; 328 C; 335 G; 381 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 6.21e-58 Length: 1541
XX Score: 627.00 Matches: 110
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 3 Gaps: 0
XX
XX US-09-940-235-4_COPY_150_259 (1-110) x AAA37644 (1-1541)
```

QY	1	ProIleAlaGluLysCysPheAspHisAalaGlyThrSerTyrValValGlyGluThr	20
Db	1209	CCCATAGCTGAGAAGTGTTTGTATCATCGCTGGACTTCCTATGTGGTTCGAGAAACG	1268
QY	21	TTrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer	40
Db	1269	TGGGAGAAGCCCTACCAGAGCTGGATGATGTGAATTGTACTTGCCTGGGAGAGGCAGC	1328
QY	41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db	1329	GGAGGCATCACTTGCACCTTCAGNAATAGATGCAACGATCAGGACACAGGACATCTTAT	1388
QY	61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Db	1389	AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	1448
QY	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	1449	ACAGGCCAACCGCCGAGGAGAGTGGAACTGTGAGAGGCACACCTCTGTGCAGACCACATCG	1508
QY	101_SerGlySerGlyProPheThrAspValArg	110	
Db	1509	AGCGGATCTGGCCCCCTCACCGATGTTTCTG	1538
RESULT 3			
AAA37643			
ID	AAA37643 standard; DNA; 2096 BP.		
XX			
AC	AAA37643;		
XX			
DT	15-SEP-2003 (revised)		
DT	13-OCT-2000 (first entry)		
XX			
DE	Chimeric SK-FBD coding sequence.		
XX			
KW	Streptokinase; SK; hybrid plasminogen activator; fibrin binding region;		
KW	plasminogen; human; fibronectin; thrombolytic therapy;		
KW	cardiovascular disorder; ss.		
XX			
OS	Streptococcus dysgalactiae subsp. equisimilis.		
OS	Homo sapiens.		
OS	Chimeric.		
XX			
PN	EP1024192-A2.		
XX			
PD	02-AUG-2000.		
XX			
Pf	23-DEC-1999; 99EP-00310541.		
XX			
PR	24-DEC-1998; 98IN-DE003825.		
XX			
PA	(COUL) CSIR COUNCIL SCI IND RES.		
XX			
PI	Sahni G, Kumar R, Roy C, Rajogopal K, Nihalani D, Sundaram V;		
PI	Yadav M;		
DR	WPI; 2000-516032/47.		
XX			
PT	Hybrid streptokinase-fibrin binding domain polypeptides useful for		
PT	thrombolytic therapy comprises a streptokinase fused with fibrin binding		
XX	domains of human fibronectin.		
PS	Example 6; Fig 22b; 58pp; English.		
XX			
CC	This sequence represents a chimeric streptokinase-fibrin binding domain		
CC	(SK-FBD) protein coding sequence. The invention relates to a hybrid		
CC	plasminogen activator (PA) comprises a polypeptide fusion between		
CC	streptokinase (SK), which are capable of plasminogen (PG) activation, and		
CC	fibrin binding regions of human fibronectin, which are from fibrin		
CC	binding domains (FBD) 4 and 5 or 1 and 2. The hybrid PA possesses the		
CC	ability to bind with fibrin independently and also characteristically		
CC	retains a PG activation ability which becomes evident only after a		

CC	pronounced duration, or lag, after exposure of the PA to a suitable		
CC	animal or human PG. The hybrid streptokinase-fibrin binding domain		
CC	polypeptides are useful in thrombolytic therapy for various kinds of		
CC	cardiovascular disorders. The hybrids have enhanced fibrin selectivity as		
CC	well as kinetics of plasminogen activation that are distinct from that of		
CC	natural streptokinase in being characterised by a temporary delay, or lag		
CC	of several minutes in the natural rate of the catalytic conversion of		
CC	plasminogen to plasmin (i.e. delayed-action thrombolysis). The proteins		
CC	can bind tightly with fibrin in blood clots soon after introduction into		
CC	the vascular system without significantly activating the circulating		
CC	blood plasminogen to plasmin, thus aiding in the localisation of the		
CC	plasminogen activation process to the site of pathological thrombus. This		
CC	overcomes systemic plasminogen activation encountered during clinical use		
CC	of streptokinase. (Updated on 15-SEP-2003 to standardise OS field)		
XX			
SQ	Sequence# 2096 BP; 643 A; 480 C; 484 G; 489 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	9.13e-58	Length:	2096
Score:	627.00	Matches:	110
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0
US-09-940-235-4_COPY_150_259 (1-110) x AAA37643 (1-2096)			
QY	1	ProIleAlaGluLysCysPheAspHisAalaGlyThrSerTyrValValGlyGluThr	20
Db	1764	CCCATAGCTGAGAAGTGTTTGTATCATCGCTGGGACTTCCTATGTGGTTCGAGAAACG	1823
QY	21	TTrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer	40
Db	1824	TGGGAGAAGCCCTACCAGGCTGGATGATGTGTAGATTGTACTTGCCTGGGAGAGGCAGC	1883
QY	41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db	1884	GGACGCATCACTTGCACCTTCAGAAATAGATGCAACGATCAGGACACAAGGACATCTAT	1943
QY	61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Db	1944	AGAAITGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC	2003
QY	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	2004	ACAGGCCAACCGCCGAGGAGAGTGGAACTGTGAGAGGCACACCTCTGTGCAGACCACATCG	2063
QY	101	SerGlySerGlyProPheThrAspValArg	110
Db	2064	AGCGGATCTGGCCCCCTCACCGATGTTTCTG	2093
RESULT 4			
ADD18477			
ID	ADD18477 standard; DNA; 2127 BP.		
XX			
AC	ADD18477;		
XX			
DT	15-JAN-2004 (first entry)		
XX			
DE	Human prostate cancer diagnosis related DNA sequence SeqID49.		
XX			
KW	prostate tissue; cancer diagnostic; cancer marker; prostate cancer; PCA;		
KW	male cancer-related death; serum biomarker; tissue biomarker; cytostatic;		
KW	gene therapy; prostate biopsy tissue; AMACR;		
KW	alpha-methylacyl-coenzyme A racemase; diagnosing cancer; cell growth;		
KW	human; ds.		
OS	Homo sapiens.		
XX			
PN	WO2003012067-A2.		
XX			

PF 02-AUG-2002; 2002WO-US024567.
XX
XX 02-AUG-2001; 2001US-0309581P.
PR 15-NOV-2001; 2001US-0334468P.
PR 01-AUG-2002; 2002US-00210120.
XX
XX (UNMI) UNIV MICHIGAN.
XX
XX Rubin MA, Chinnaiyan AM, Sreekumar A;
XX WPI; 2003-278396/27.
XX
XX Characterizing prostate tissue comprises providing a prostate tissue
XX sample from a subject and detecting the presence or absence of expression
XX of hepsin, pim-1 or EZH2.
XX
XX Disclosure; SEQ ID NO 49; 297pp; English.
XX
XX This invention relates to a novel method of characterising prostate
XX tissue in a subject and to compositions and methods for cancer
XX diagnostics, including cancer markers, in particular prostate cancer.
XX Prostate cancer (PCA) is a leading cause of male cancer-related death.
XX Additional serum and tissue biomarkers would aid diagnosis. The invention
XX may provide means of producing compounds with a cytostatic activity or
XX allow the development of gene therapy. The methods of the invention
XX useful for characterising prostate tissue in a subject, screening
XX compounds, characterising inconclusive prostate biopsy tissue in a
XX subject, detecting AMACR (alpha-methylacyl-coenzyme A racemase)
XX expression in a bodily fluid, characterising tissue in a subject,
XX diagnosing cancer in a subject and inhibiting the growth of cells. The
XX present sequence is a DNA sequence which is preferably utilised in the
XX method of the invention.
XX
XX SQ Sequence 2127 BP; 552 A; 511 C; 576 G; 488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9, 3e-58 Length: 2127
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ADD18477 (1-2127)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 597 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 656

QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
DB 657 TGGGAGAACCCCTACCAAGCTGGAGTATGATGGTAGATTGTACTTGCCTGGGAGAAGCAGC 716

QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 717 GGACCATCATCTTGCACTTCTAGAAATAGATGATCAACGATCAGACACAAAGGACATCCTAT 776

QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
DB 777 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 836

QY 81 ThrGlyAsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrSer 100
DB 837 ACAGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 896

QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 897 AGCGGATCTGGCCCTTCACCGATGTTCTG 926

RESULT 5
AAx81299 standard; cDNA; 2147 BP.
ID
XX

AC AAX81299;
XX
XX 21-SEP-1999 (first entry)
XX
XX Human migration stimulating factor (MSF) 1-alpha encoding cDNA.
DE
XX Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha; ss.
XX
XX Homo sapiens.
OS
XX WO9931233-A1.
PN
XX 24-JUN-1999.
PD
XX 15-DEC-1998; 98WO-GB003766.
PF
XX 16-DEC-1997; 97GB-00026539.
PR
XX (UYDU-) UNIV DUNDEE.
PA
XX Schor SL, Schor AM;
PI
XX WPI; 1999-430039/36.
XX
XX P-PSDB; AAV28901.
DR
XX Proteins with cell migration stimulatory activity used in treating wound
XX and preventing scarring.
PT
XX Example 1; Fig 1; 86pp; English.
PS
XX The invention provides a human migration stimulatory factor (MSF)
XX protein. Host cells containing a replicable vector comprising the MSF
XX encoding nucleic acid can be used for the recombinant production of the
XX protein. The polypeptide can be used for modulating cell migration,
XX healing a wound and for preventing scarring. The present sequence
XX represents the nucleotide sequence encoding a human MSF1-alpha protein
XX
XX SQ Sequence 2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9, 41e-58 Length: 2147
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x AAX81299 (1-2147)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 597 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 656

QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
DB 657 TGGGAGAACCCCTACCAAGCTGGAGTATGATGGTAGATTGTACTTGCCTGGGAGAAGCAGC 716

QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 717 GGACCATCATCTTGCACTTCTAGAAATAGATGATCAACGATCAGACACAAAGGACATCCTAT 776

QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
DB 777 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 836

QY 81 ThrGlyAsnGlyArgGlyGluThrLysCysGluArgHisThrSerValGlnThrSer 100
DB 837 ACAGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCAATCG 896

QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 897 AGCGGATCTGGCCCTTCACCGATGTTCTG 926

```
RESULT 6
ADP75954
ID ADP75954 standard; DNA; 3540 BP.
XX
AC ADP75954;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human mini fibronectin gene SeqID4.
XX
KW matrix binding region; cell surface; extracellular matrix;
KW leukaemia inhibitory factor region; cell differentiation;
KW embryonic stem cell; gene; ds; human; mini fibronectin.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3540
FT /tag= a
FT /product= "Human mini fibronectin protein"
FT /partial
FT /note= "No start codon"
XX
PN JP2004166641-A.
XX
XX 17-JUN-2004.
XX
XX 21-NOV-2002; 2002JP-00338373.
XX
XX 21-NOV-2002; 2002JP-00338373.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX
XX (AIZU/) AIZU Y.
XX
XX WPI; 2004-445579/42.
XX
XX P-PSDB; ADP75952.
XX
XX Novel protein which has matrix binding region which assembles to cell
XX surface, or to extracellular matrix surrounding cell and leukemia
XX inhibitory factor region that suppresses cell differentiation of
XX embryonic stem cell.
XX
XX Example 1; SEQ ID NO 4; 49pp; Japanese.
XX
XX This invention relates to a novel protein (and the gene which encodes it)
XX which has a matrix binding region which assembles to a cell surface, or
XX extracellular matrix surrounding the cell, and a leukaemia inhibitory
XX factor region which suppresses cell differentiation of embryonic stem
XX cells. The invention is useful for suppressing cell differentiation. The
XX invention can be used for the efficient inhibition of embryonic stem cell
XX differentiation. The invention thus enables efficient and continuous
XX proliferation of embryonic stem cells. Therefore, embryonic stem cells
XX can be prepared in large quantities and utilised for fundamental research
XX and applications. The present sequence is that of a gene which encodes
XX the human mini-fibronectin protein which was used in the exemplification
XX of the invention.
XX
XX SQ Sequence 3540 BP; 952 A; 893 C; 928 G; 767 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,768-57 Length: 3540
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
XX
US-09-940-235-4_COPY_150_259 (1-110) x ADP75954 (1-3540)
XX
Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 448 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGCTGGAGAAACG 507
```

```
21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
508 TGGGAGAGAGCCCTACCAAGGCTGGATGATGGTAGATTGTTCTTGGCTGGGAGAGGAGC 567
41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
568 GGACGCATCATCTTGCACCTTCTAGAAATAGATGATCAACGATCAGGACACAAAGACATCTAT 627
61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
628 AGAATTGGAGACACCTTGGACCAAGAGGATATTCGAGAAACCTGCTCCAGTGCATCTGC 687
81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
688 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG 747
101 SerGlySerGlyProPheThrAspValArg 110
748 AGCGGATCTGGCCCTTCACCGATGTTCTG 777
RESULT 7
ADP75958
ID ADP75958 standard; DNA; 4080 BP.
XX
AC ADP75958;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human leukaemia inhibiting factor/mini fibronectin gene SeqID8.
XX
KW matrix binding region; cell surface; extracellular matrix;
KW leukaemia inhibitory factor region; cell differentiation;
KW embryonic stem cell; gene; ds; human; mini fibronectin.
XX
OS Homo sapiens.
XX
XX Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..4080
FT /tag= a
FT /product= "Human leukaemia inhibiting factor/mini
FT fibronectin chimeric protein"
FT /partial
FT /note= "No start codon"
XX
XX JP2004166641-A.
XX
XX 17-JUN-2004.
XX
XX 21-NOV-2002; 2002JP-00338373.
XX
XX 21-NOV-2002; 2002JP-00338373.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN KAGAKU GIJUTSU SH.
XX
XX (AIZU/) AIZU Y.
XX
XX WPI; 2004-445579/42.
XX
XX P-PSDB; ADP75957.
XX
XX Novel protein which has matrix binding region which assembles to cell
XX surface, or to extracellular matrix surrounding cell and leukemia
XX inhibitory factor region that suppresses cell differentiation of
XX embryonic stem cell.
XX
XX Disclosure; SEQ ID NO 8; 49pp; Japanese.
XX
XX This invention relates to a novel protein (and the gene which encodes it)
XX which has a matrix binding region which assembles to a cell surface, or
XX extracellular matrix surrounding the cell, and a leukaemia inhibitory
XX factor region which suppresses cell differentiation of embryonic stem
XX cells. The invention is useful for suppressing cell differentiation. The
XX invention can be used for the efficient inhibition of embryonic stem cell
```

CC differentiation. The invention thus enables efficient and continuous
CC proliferation of embryonic stem cells. Therefore, embryonic stem cells
CC can be prepared in large quantities and utilised for fundamental research
CC and applications. The present sequence is that of the DNA sequence which
CC encodes the protein of the invention.

SQ Sequence 4080 BP; 1075 A; 1086 C; 1058 G; 861 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.1e-57 Length: 4080
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ADP75958 (1-4080)

Qy 1 ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 448 CCCATAGCTGAGAAGTGTTCATGCTGCTGGGACTTCCTATGTGTGGAGAAACG 507
Qy 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGluGlySer 40
Db 508 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCAGC 567
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 568 GGACGCATCACTTGCACCTCTAGAAATAGATGATCAACGATCAGGACACAGGACATCTAT 627
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 628 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 687
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 688 ACAGGCAAGCGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCATCG 747
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 748 AGCGGATCTGGCCCTTCACCGATGTTGCT 777

RESULT 8

ACC72037
ID ACC72037 standard; DNA; 4295 BP.

XX AC ACC72037;
XX DT 08-JUL-2003 (first entry)

XX DE BC00770 gene #SEQ ID 51.
XX KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
XX KW drug discovery; clinical medicine; forensic medicine; gene;
XX KW chromosome 2q35; ds.

XX OS Homo sapiens.
XX PN WO2003029421-A2.
XX PD 10-APR-2003.

XX XX 02-OCT-2002; 2002WO-US031287.
XX PF 03-OCT-2001; 2001US-0326526P.
XX PR 14-MAY-2002; 2002US-00144194.

XX XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX XX WPI; 2003-381623/36.
XX DR P-PSDB; ABR58303.

XX
PT
PT
PT
XX
XX
XX

New isolated human differentially-regulated breast cancer polynucleotide
and polypeptide, useful for diagnosing, staging, prognosticating,
preventing and/or treating diseases and conditions relating to breast
cancer.

Claim 2; SEQ ID NO 51; 127pp + Sequence Listing; English.

The invention relates to isolated polynucleotides which are
differentially-regulated in breast cancer. The methods and compositions
of the present invention are useful for detecting, diagnosing, staging,
monitoring, prognosticating, preventing and/or treating diseases and
conditions relating to breast cancer, and may be used in gene therapy or
antisense therapy. They can also be used in research, drug discovery,
clinical medicine and forensic medicine. Sequences given in records,
ACC72012-ACC72074 represent polynucleotides of the invention that are
differentially-regulated in breast cancer. NOTE: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 4295 BP; 1287 A; 866 C; 1066 G; 1076 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.24e-57 Length: 4295
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ACC72037 (1-4295)

Qy 1 ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 867 CCCATAGCTGAGAAGTGTTCATGCTGCTGGGACTTCCTATGTGTGGAGAAACG 926
Qy 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGluGlySer 40
Db 927 TGGGAGAAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGCCTGGGAGAGGCAGC 986
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 987 GGACGCATCACTTGCACCTCTAGAAATAGATGACAGATCAGGACACAGGACATCTAT 1046
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 1047 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 1106
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 1107 ACAGGCAAGCGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGAGACCATCG 1166
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 1167 AGCGGATCTGGCCCTTCACCGATGTTGCT 1196

RESULT 9

AAA35009
ID AAA35009 standard; DNA; 4860 BP.

XX AC AAA35009;
XX XX 28-JUL-2000 (first entry)

XX DT Human adenosine receptor related polynucleotide SEQ ID NO:2698.
XX DE Human, adenosine receptor; low adenosine antisense oligonucleotide;
XX KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;

PA (PADI/) PADIGARU M.
PA (GUSE/) GUSEV V Y.
PA (MALY/) MALYANKAR U M.
PA (BURG/) BURGESS C E.
PA (GERL/) GERLACH V.
PA (CASM/) CASMAN S J.
PA (RIEG/) RIEGER D K.
PA (GROS/) GROSSE W M.
PA (SMIT/) SMITHSON G.
PA (PEYM/) PEYMAN J A.
PA (STAR/) STARLING G.
PA (ROTH/) ROTHENBERG M E.
PA (LARO/) LAROCHELLE W J.
PA (SHIM/) SHIMKETS R A.
PA (CRAB/) CRABTREE J.
PA (RAST/) RASTELLI L.
PA (VOSS/) VOSS E Z.
PA (BOLD/) BOLDG F L.
PA (EDIN/) EDINGER S R.
PA (MILL/) MILLET I.
PA (MACD/) MACDOUGALL J R.
PA (ELLE/) ELLERMAN K.
PA (CHAP/) CHAPOVAL A.
XX
PI Zhong M, Li L, Gorman L, Spytek KA, Kekuda R, Taupier RJ;
PI Anderson DW, Vernet CAM, Catterton E, Miller CE, Shenoy SG;
PI Patturajan M, Pena CE, Tchernev VT, Padigar M, Gusev VY;
PI Malyankar UM, Burgess CE, Gerlach V, Casman SJ, Rieger DK;
PI Grosse WM, Smithson G, Peyman JA, Starling G, Rothenberg ME;
PI Larochelle WJ, Shimkets RA, Crabtree J, Rastelli L, Voss EZ;
PI Boldog FL, Edinger SR, Millet I, Macdougall JR, Ellerman K;
PI Chapoval A;
XX
DR WPI: 2004-355290/33.
DR P-PSDB; ADN95948.
XX
PT New isolated polypeptide, useful for treating or preventing a pathology
PT associated with the polypeptide, e.g. diabetes, infectious disease,
PT cancer, neurodegenerative disorders or Alzheimer's disease.
XX
PS Claim 20; SEQ ID NO 1; 552pp; English.
XX
CC The invention relates to human NOVX polypeptides and polynucleotides. The
CC isolated nucleic acids can be used to express the novel proteins, to
CC detect novel mRNA or a genetic lesion in a novel gene and to modulate its
CC activity. It can also be used in gene therapy for treating or preventing
CC a pathology associated with the protein or nucleic acid. The disorders
CC include metabolic disorders, diabetes, obesity, infectious diseases,
CC anorexia, cancer, neurodegenerative disorders, Alzheimer's disease,
CC Parkinson's disease, immune disorders and haematopoietic disorders. This
CC sequence represents a human NOVX polynucleotide of the invention.
XX
SQ Sequence 7049 BP; 1888 A; 1972 C; 1651 G; 1535 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 4,17e-57 Length: 7049
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ADN95947 (1-7049)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 461 CCCATAGCTGAGAGTGTGTTGATCATGCTGCTGGGACTTCCTATGCTGGGAGAAACG 520
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
Db 521 TGGGAGAGCCCTACCAAGGCTGGATGATGGTATGTTACTTGCCTGGGAGAGGAGC 580
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60

Db 581 GGACGCATCACTTGTCACTTCTAGAAATAGATGCAACGATCAGGACACAAGGACATCCTAT 640
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGAGCAAGAAGATATTCAGGAAACCTGCTCCAGTGCATCTGC 700
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGGCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG 760
QY 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGCGCCCTTCACCGATGTTTCGT 790

RESULT 12
ACD06170
ID ACD06170 standard; cDNA; 7361 BP.
XX
AC ACD06170;
XX
DT 06-AUG-2003 (first entry)
XX
DE Human cDNA encoding protein NOV1b.
XX
KW Human; ss; gene; NOVX; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; prostate cancer; diabetes; metabolic disorder;
KW neoplasm; graft versus host disease; AIDS; bronchial asthma;
KW Crohn's disease; multiple sclerosis; infectious disease; anorexia;
KW cancer-associated cachexia; neurodegenerative disorder; snp;
KW Alzheimer's disease; Parkinson's disease; immune disorder;
KW haematopoietic disorder; dyslipidaemia; wasting disorder; gene therapy;
KW single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation /tag= a replace(1921,C)
FT /standard_name= "Single nucleotide polymorphism"
FT variation /*tag= b replace(4730,G)
FT /standard_name= "Single nucleotide polymorphism"
FT variation /tag= c replace(6395,C)
FT /standard_name= "Single nucleotide polymorphism"
XX
PN WO2003023008-A2.
XX
PD 20-MAR-2003.
XX
PF 09-SEP-2002; 2002WO-US028596.
XX
PR 07-SEP-2001; 2001US-0318120P.
PR 07-SEP-2001; 2001US-0318130P.
PR 10-SEP-2001; 2001US-0318430P.
PR 12-SEP-2001; 2001US-0318765P.
PR 17-SEP-2001; 2001US-0322781P.
PR 17-SEP-2001; 2001US-0322816P.
PR 19-SEP-2001; 2001US-0323519P.
PR 20-SEP-2001; 2001US-0323631P.
PR 20-SEP-2001; 2001US-0323636P.
PR 25-SEP-2001; 2001US-0324969P.
PR 25-SEP-2001; 2001US-0325091P.
PR 26-SEP-2001; 2001US-0324990P.
PR 15-FEB-2002; 2002US-0357303P.
PR 28-FEB-2002; 2002US-0360973P.
PR 20-MAR-2002; 2002US-0366131P.
PR 25-MAR-2002; 2002US-0367753P.
PR 02-APR-2002; 2002US-0369479P.
PR 10-MAY-2002; 2002US-0379532P.
PR 17-MAY-2002; 2002US-0381664P.
PR 17-MAY-2002; 2002US-0381672P.

PR 28-MAY-2002; 2002US-0383651P.
PR 19-JUN-2002; 2002US-0384012P.
PR 29-MAY-2002; 2002US-0390155P.
PR 06-SEP-2002; 2002US-00390155.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Zhong M, Li L, Gorman L, Spytek KA, Kekuda R, Taupier RJ;
PI Anderson DW, Vernet CAM, Catterton E, Miller CE, Shenoy SG;
PI Paturajan M, Pena CE, Tchernev VT, Padigar M, Gusev VV;
PI Malyankar UM, Burgess CE, Gerlach VL, Casman SJ, Rieger DK;
PI Grosse WM, Smithson G, Peyman JA, Starling G, Rothenberg ME;
PI Larochelelle WJ, Shinkets RA, Crabtree J, Rastelli L, Voss EZ;
PI Boldog FL, Edinger SR, Millet I, Macdougall JR, Ellerman K;
PI Chapoval A;
XX
XX WPI; 2003-313246/30.
DR P-PSDB; ABO01289.
DR
XX
XX New polypeptides and polynucleotides having properties related to
PT stimulation of biochemical or physiological responses in a cell or
PT tissue, useful for diagnosing or preventing e.g. atherosclerosis,
PT hypertension, prostate cancer.
XX
XX Claim 20; Page 108-110; 849pp; English.
XX
XX The invention relates to an isolated polypeptide comprising one of 127
CC sequences (appearing as ABO1288-ABO1414) designated as NOVX, a mature
CC form of NOVX, an amino acid sequence which is at least 95% identical to
CC NOVX or an amino acid sequence comprising one or more conservative
CC substitutions in NOVX. Also included are nucleic acids encoding NOVX
CC proteins, determining the presence or amount of NOVX or NOVX DNA in a
CC sample (by introducing the sample to an antibody that binds
CC immunospecifically to the polypeptide, and determining the presence or
CC amount of antibody bound to the polypeptide), determining the presence of
CC or predisposition to a disease associated with altered levels of
CC expression of NOVX or NOVX DNA in a first mammalian subject, identifying
CC an agent that binds to NOVX, identifying a potential therapeutic agent
CC for treatment of a pathology related to aberrant expression or aberrant
CC physiological interactions of NOVX, screening for a modulator of activity
CC of or of latency or predisposition to a pathology associated with NOVX, a
CC vector comprising NOVX DNA, a cell comprising the vector (used to produce
CC NOVX) and an anti-NOVX antibody. The NOVX nucleic acids and polypeptides
CC are useful as a marker for cell or tissue type, and in diagnosing and
CC treating pathologies, diseases, conditions or disorders associated with
CC NOVX sequences, including cardiomyopathy, atherosclerosis, hypertension,
CC congenital heart defects, prostate cancer, diabetes, metabolic disorders,
CC neoplasm, graft versus host disease, AIDS, bronchial asthma, Crohn's
CC disease, multiple sclerosis, infectious diseases, anorexia, cancer-
CC associated cachexia, neurodegenerative disorders (e.g. Alzheimer's
CC disease or Parkinson's disease), immune disorders, haematopoietic
CC disorders, dyslipidaemias, and wasting disorders associated with chronic
CC diseases. These may also be used to screen for molecules which inhibit or
CC enhance NOVX activity or function, and for detecting specific cell types.
CC These may also be used in chromosome mapping, gene therapy, tissue
CC typing, and in forensic biology. The present sequence encodes a NOVX
CC protein
XX
SQ Sequence 7361 BP; 1986 A; 1892 C; 1806 G; 1673 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 4,41e-57 Length: 7361
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ACD06170 (1-7361)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 546 CCCATAGTGTAGAGTGTTCATGTCATGCTGGGACCTCTCTGTCGGAGGACG 605

Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
Db 606 TGGGAGAGCCCTACCAAGCTGGATGATGTAGATTGCTTGGCTGGGAGAGCAGC 665
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 666 GGACGCATCATTGCACTTCTAGAAATAGATGATCAACGATCAGGACACCAAGCATCTAT 725
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 726 AGAATTGGAGACACCTGGAGCAAGAGGATATCGAGGAAACCTGCTCCAGTGCATCTGC 785
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 786 ACAGGCACGCGGAGGAGTGAAGTGTGAGAGGACACACCTCTGTGCAGACCATCG 845
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 846 ACGCGATCTGGCCCTTCACCGATGTTGCT 875
RESULT 13
ACC00412
ID ACC00412 standard; cDNA; 7550 BP.
XX
XX ACC00412;
XX
XX 04-JUL-2003 (first entry)
DT
DE Human cell adhesion and extracellular matrix protein, CADECM-21, DNA.
XX
XX Human; anti-HIV; antiallergic; cerebroprotective; antiparkinsonian;
KW anticonvulsant; nootropic; neuroprotective; immunosuppressive;
KW dermatological; anti-inflammatory; cytostatic; antiarteriosclerotic;
KW gene therapy; cell adhesion; extracellular matrix; CADSCM;
KW immune system disorder; AIDS; allergy; neurological disorder; stroke;
KW Parkinson's disease; epilepsy; developmental disorder; Down's syndrome;
KW cerebral palsy; connective tissue disorder; systemic lupus erythematosus;
KW genetic disorder; Alport's syndrome; cell proliferative disorder; cancer;
KW atherosclerosis; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 271..7071
FT FT /*tag= a
FT FT /product= "CADECM-21"
XX
XX WO2003027230-A2.
XX
XX 03-APR-2003.
XX
XX 02-AUG-2002; 2002WO-US024649.
XX
XX 03-AUG-2001; 2001US-0309564P.
XX 03-AUG-2001; 2001US-0310119P.
XX 17-AUG-2001; 2001US-0313091P.
XX 31-AUG-2001; 2001US-0316771P.
XX 07-SEP-2001; 2001US-0317896P.
XX 21-SEP-2001; 2001US-0324781P.
XX 05-OCT-2001; 2001US-0327606P.
XX 12-OCT-2001; 2001US-0328960P.
XX 09-NOV-2001; 2001US-034471P.
XX 17-MAY-2002; 2002US-0381291P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Burford N, Warren BA, Duggan BM, Mason PM, Richardson TW, Yue H;
PI Forsythe IJ, Elliott VS, Griffin JA, Gorvad AB, Azimzai Y;
PI Kallick DA, Xu Y, Honchell CD, Baughn MR, Gietzen KJ, Lee S;
PI Walla NK, Tang YT, Nguyen DB, Becha SD, Lee SY, Ramkumar J;
XX
XX WPI; 2003-354645/33.
DR

DR P-PSDB; ABR40124.
XX
XX New human cell adhesion and extracellular matrix proteins (CADECM),
PT useful for diagnosing, treating or preventing disorders associated with
PT aberrant CADECM expression, e.g. cancer, AIDS, atherosclerosis, allergies
PT or stroke.
XX
XX Claim 12; Page 231-233; 234pp; English.
XX
XX The present invention relates to novel human cell adhesion and
CC extracellular matrix proteins, CADECM-1 to CADECM-22, and their coding
CC sequences (ACC00392-ACC00413 and ABR40104-ABR40125). The coding sequences
CC and proteins are useful in diagnosing, treating and preventing disorders
CC associated with aberrant expression of CADECM, such as immune system
CC disorders (e.g. AIDS or allergies), neurological disorders (e.g. stroke,
CC Parkinson's disease or epilepsy), developmental disorders (e.g. Down's
CC syndrome or cerebral palsy), connective tissue disorders (e.g. systemic
CC lupus erythematosus), genetic disorders (e.g. Alport's syndrome) or cell
CC proliferative disorders (e.g. cancer or atherosclerosis)
XX
SQ Sequence 7550 BP; 1996 A; 1993 C; 1889 G; 1672 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,55e-57 Length: 7550
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ACC00412 (1-7550)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 811 CCCATAGCTGAGAAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGAGAACG 870
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
DB 871 TGGGAGAGCCCTACCAAGCTGGATGATGTTGATGTTGCTTGGGAGAGGACG 930
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 931 GGACGCATCAGTTCGACTTCTAGAAATAGATGATGCAACGATCAGGACACCAAGGACATCTAT 990
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 991 AGAATGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 1050
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 1051 ACAGGCAACGGCCGAGGAGTGGAACTGTGAGAGGCACACCTCTGTGCAGACCATCG 1110
QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 1111 AGCGGATCTGGCCCTTCACCGATGTTGCT 1140

RESULT 14

ADB31322

ID ADB31322 standard; cDNA; 7679 BP.

XX AC

XX ADB31322;

XX 04-DEC-2003 (first entry)

XX DT

XX Testosterone regulated prostate cancer gene #38.

XX DE

XX Human; ss; prostate cancer; cancer; androgen; castration;

XX KW

XX anti-androgenic drug; bicalutamide; casodex; testosterone.

XX OS

XX Homo sapiens.

XX PN

XX US6506607-B1.

PD

XX 14-JAN-2003.

XX

PF 23-DEC-1998; 98US-00220132.

XX

PR 24-DEC-1997; 97US-0068821P.

XX

PR 25-MAR-1998; 98US-0079303P.

XX

XX (MILL-) MILLENNIUM PHARM INC.

XX

XX Shyjan AW;

XX

XX WPI; 2003-595981/56.

XX

XX Determination of whether compound is useful for prostate cancer treatment

XX comprises measuring expression level of specific nucleic acid sequence in

XX prostate cancer cell sample in the presence and absence of compound.

XX

XX Example 1; Col 139-146; 194pp; English.

XX

XX The invention discloses a method for determining whether a compound can
CC be used to treat prostate cancer and comprises measuring the expression
CC level of a nucleic acid in prostate cancer cell sample in the presence
CC and absence of the compound. Determining whether a compound can be used
CC to treat prostate cancer comprises identifying the compound as useful for
CC prostate cancer treatment when the expression level of the nucleic acid
CC in the presence of the compound is less than the expression level of the
CC nucleic acid in the absence of the compound. Prostate cancer is usually
CC treated by androgen withdrawal, by castration or through the use of an
CC anti-androgenic drug. Bicalutamide (casodex) is one such anti-androgenic
CC compound. The determination whether a selected compound, e.g. an anti-
CC androgenic compound or testosterone, can be used to treat the prostate
CC cancer can be made on a patient by patient basis. The sequence presented
CC is a gene which is more highly expressed in testosterone treated prostate
CC cancer cells than the untreated cells.

SQ Sequence 7679 BP; 2081 A; 1981 C; 1862 G; 1755 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,65e-57 Length: 7679
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x ADB31322 (1-7679)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 461 CCCATAGCTGAGAAGTGTGTTTGTATCATGCTGCTGGGACTTCTTATGTTGGTGGAGAACG 520
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
DB 521 TGGGAGAGCCCTACCAAGCTGGATGATGTTGATGTTGCTTGGGAGAGGACG 580
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 581 GGACGCATCAGTTCGACTTCTAGAAATAGATGATGCAACGATCAGGACACCAAGGACATCTAT 640
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 641 AGAATGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 700
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 701 ACAGGCAACGGCCGAGGAGTGGAACTGTGAGAGGCACACCTCTGTGCAGACCATCG 760
QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 761 AGCGGATCTGGCCCTTCACCGATGTTGCT 790

RESULT 15

AAT17551

ID AAT17551 standard; cDNA; 7680 BP.
XX AC AAT17551;
XX DT 21-JUN-1996 (first entry)
XX DE Human fibronectin cDNA.
XX KW Fibrin-binding peptide; fibronectin; fibrinolysis; fibrinogenesis;
KW thrombus; pulmonary embolus; atherosclerosis; tumour; diagnosis; therapy;
KW ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 14..6988
FT /*tag= a
XX PN WO9604304-A1.
XX PD 15-FEB-1996.
XX PF 01-AUG-1995; 95WO-US009819.
XX PR 01-AUG-1994; 94US-00283857.
XX PA (UNYNY) UNIV NEW YORK STATE.
XX PI Gold LI, Rostagno AA;
XX WPI; 1996-129333/13.
XX DR P-PSDB; AAR92778.
XX PT New fibrin-binding peptide molecules - used for the diagnosis and
PT treatment of conditions associated with fibrin deposition, e.g. thrombi.
XX PS Disclosure; Page 112-115; 146pp; English.
XX CC A cDNA sequence (AAT17551) codes for human fibronectin (AAR92778). It is
CC used for the prodn. of fibrin-binding modules of fibronectin useful in
CC the diagnosis and treatment of conditions associated with fibrin
CC deposition, e.g. for disruption of blood clots
XX SQ Sequence 7680 BP; 2083 A; 1979 C; 1862 G; 1756 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.65e-57 Length: 7680
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x AAT17551 (1-7680)

QY 1 ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrrValValGlyGluThr 20
DB 461 CCCATAGCTGAGAAGTGTTCATCATGCTGTGGACTTCCTATGTCGGAGAAACG 520
QY 21 TrpGluLysProTyrrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
DB 521 TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGTTACTTCCTGGAGAGGCAGC 580
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrr 60
DB 581 GGACGCATCCTTGCACCTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 640
QY 61 ArgIleGlyAspThrTrpSerLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
DB 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 700
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100

Db 701 ACAGGCAACGGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTTCTGTGCAGACCACATCG 760
QY 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCCCTTCACCCGAATTCGT 790

Search completed: November 6, 2004, 00:38:01
Job time : 184.113 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2004, 00:01:06 ; Search time 1343.92 Seconds
(without alignments)
2982.593 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAEKCFDHAAGTSYVVGET.....ERHSTVQTSSGSGPFTDVR 110

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV=xlp

-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_03112004_174039_11264/app_query.fasta_1.1045
-DB=EST -QFMT=FASTAP -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09940235@cgn_1_1_8076@runat_03112004_174039_11264 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*\n2: gb_est2:*\n3: gb_hic:*\n4: gb_est3:*\n5: gb_est4:*\n6: gb_est5:*\n7: gb_est6:*\n8: gb_gss1:*\n9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	627	100.0	451	AL603368	AL603368 DKFZp686C
2	627	100.0	560	CN419585	CN419585 170005318
3	627	100.0	621	CN482442	CN482442 hw200808.y
4	627	100.0	626	CN419611	CN419611 170005313
5	627	100.0	737	AU140993	AU140993 AU140993
6	627	100.0	765	BM715855	BM715855 UI-E-EJO-
7	627	100.0	861	AU140971	AU140971 AU140971
8	627	100.0	7501	BC078656	BC078656 Homo sapi
9	620	98.9	734	AU140802	AU140802 AU140802

c	10	614	97.9	465	1	AI095589	AI095589 qb24a08.x
c	11	614	97.9	500	2	BF956982	BF956982 QVA-NN114
c	12	610	97.3	560	1	AI743013	AI743013 wg85a12.x
c	13	605	96.5	548	6	CD613789	CD613789 56022208J
	14	605	96.5	551	6	CD613788	CD613788 56022208H
	15	602	96.0	547	2	BE009640	BE009640 PM4-EN017
	16	601	95.9	603	7	CN419658	CN419658 170005315
c	17	600	95.7	943	5	BX391752	BX391752 BX391752
	18	597	95.1	705	1	AU140789	AU140789 AU140789
	19	590	94.2	474	6	CB536952	CB536952 772021 MA
	20	589	93.9	739	1	AL706215	AL706215 DKFZp686J
	21	586	93.5	414	1	AA492032	AA492032 ng55a12.s
c	22	584	93.1	518	1	AI093548	AI093548 qb08h09.x
	23	581.5	92.7	707	7	CK639246	CK639246 UI-M-HSO-
	24	581.5	92.7	836	7	CN538822	CN538822 UI-M-HSO-
	25	581.5	92.7	839	7	CK129682	CK129682 AGENCOURT
	26	581.5	92.7	8315	3	AK090135	AK090135 Mus muscu
	27	581.5	92.7	8329	3	AK090130	AK090130 Mus muscu
	28	575.5	91.8	781	7	CN534124	CN534124 UI-M-HSO-
	29	570	90.9	538	1	AL603599	AL603599 DKFZp686K
	30	558	89.0	303	1	AA852090	AA852090 NHTBCae10
	31	557	88.8	496	4	BG945197	BG945197 PMO-AN008
	32	551.5	88.0	2107	3	AK054456	AK054456 Mus muscu
	33	546.5	87.2	859	7	CF616056	CF616056 AGENCOURT
	34	544.5	86.8	417	8	CC200505	CC200505 RST455 Ba
	35	543	86.6	427	7	CN419650	CN419650 170005315
	36	543	86.6	452	7	CN419649	CN419649 170005315
c	37	542.5	86.5	410	2	BE110611	BE110611 UI-R-BJ1-
	38	542	86.4	408	4	BG900107	BG900107 HOA51-1-C
	39	542	86.4	605	7	CN419657	CN419657 170004553
c	40	533	85.0	840	5	BU436156	BU436156 603220793
	41	523	83.4	495	2	BF956977	BF956977 QVA-NN114
	42	523	82.9	494	4	BF956365	BF956365 QV1-GN020
	43	517	82.5	580	5	BQ292415	BQ292415 PMO-AN008
	44	509	81.2	454	7	W46530	W46530 xc32h07.r1
	45	497	79.3	484	4	BI058354	BI058354 PM3-GN051

ALIGNMENTS

RESULT 1
AL603368
LOCUS AL603368 451 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686C067.r1.686 (synonym: hlcc3) Homo sapiens cDNA clone
AL603368 DKFZp686C067.5, mRNA sequence.
ACCESSION AL603368
VERSION AL603368.1 GI:15166874
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 451)
AUTHORS Ansoorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
TITLE EST (Ansoorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No si sequence available.
This clone (DKFZp686C067) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..451
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="DKF2p66C067"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Alignment Scores:

Pred. No.: 4,54e-59 Length: 451
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x AL603368 (1-451)

Qy 1 ProLeaAGluUyCysPheaspHisAlaAGlyThrSerTyValValGlyGluThr 20
Db 32 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGACTTCCTATGTGTGCGAGAAACG 91
Qy 21 TTPGluUysProTyTGlnGlyTTPMetMetValAspCysThrCysLeuGlyGlySer 40
Db 92 TGGAGAGAGCCCTACCAAGCTGGATGATGTAGATTGTACTTGCCTGGAGAGGACG 151
Qy 41 GlyArgileThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 60
Db 152 GGACGCATCATTGCTACCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 211
Qy 61 ArgileGlyAspThrTTPSerlysiysAspAsnArgGlyAsnleuGlnCysileCys 80
Db 212 AGAATTGGAGACACCTCGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 271
Qy 81 ThrGlyAsnGlyArgGlyGluTTPlyCysGluArgHisThrSerValGlnThrThrSer 100
Db 272 ACAGCAACGCCGAGGAGAGTGGAGTGTGAGAGCACACCTCTGTGAGACCAATCG 331
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 332 AGCGGATCTGGCCCTTCACCGATGTTCGT 361

RESULT 2

CN419585
LOCUS 17000531856084 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN419585
ACCESSION CN419585.1 GI:47407179
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 560)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Bradenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,
Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Stanton,L.W.

Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

CONTACT: Brandenberger R

Regenerative Medicine

Geront Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbradenberger@geron.com

Insert Length: 560 Std Error: 0.00.

Location/Qualifiers

1..560
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/tissue type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="oligo dT primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN

Alignment Scores:
Pred. No.: 5.99e-59 Length: 560
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x CN419585 (1-560)

Qy 1 ProLeaAGluUyCysPheaspHisAlaAGlyThrSerTyValValGlyGluThr 20
Db 209 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGACTTCCTATGTGTGCGAGAAACG 268
Qy 21 TTPGluUysProTyTGlnGlyTTPMetMetValAspCysThrCysLeuGlyGlySer 40
Db 269 TGGAGAGAGCCCTACCAAGCTGGATGATGTAGATTGTACTTGCCTGGAGAGGACG 328
Qy 41 GlyArgileThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 60
Db 329 GGACGCATCATTGCTACCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 388
Qy 61 ArgileGlyAspThrTTPSerlysiysAspAsnArgGlyAsnleuGlnCysileCys 80
Db 389 AGAATTGGAGACACCTCGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 448
Qy 81 ThrGlyAsnGlyArgGlyGluTTPlyCysGluArgHisThrSerValGlnThrThrSer 100
Db 449 ACAGCAACGCCGAGGAGAGTGGAGTGTGAGAGCACACCTCTGTGAGACCAATCG 508
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 509 AGCGGATCTGGCCCTTCACCGATGTTCGT 538

RESULT 3

CN482442
LOCUS 621 bp mRNA linear EST 26-APR-2004
DEFINITION hw20d08.y1 Human primary human ocular pericytes. Unamplified (hw)
ACCESSION Homo sapiens cDNA clone hw20d08 5', mRNA sequence.
VERSION CN482442
KEYWORDS EST.
SOURCE CN482442.1 GI:46563946
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 621)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Tsai,J.Y. and Wistow,G.
Expressed sequence tag analysis of cultured primary human ocular
pericytes

Unpublished (2004)

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331, NIH, Bethesda, MD 20892-2740, USA

Tel: 301 403 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 20 row: d column: 08

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

1..621
/organism="Homo sapiens"

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw20d08"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human primary human ocular pericytes."
Unamplified (hw)"
/note="Organ: Eye; Vector: pSPORT1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGATGTCCTAGTCGCGAGCGGCCGCTT]5-3'. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:
Pred. No.: 6,838-59 Length: 621
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x CN482442 (1-621)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 155 CCCATAGCTGAGAAGTGTTCCTGCTGGGACCTCTATGTCGGAGAACG 214
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
DB 215 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTTGCTGCTGGGAGAGCAGC 274
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 275 GGACGCATCTGCTGCTAGAAATAGATGATCAACGATCAGGACACCAAGGACATCTAT 334
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
DB 335 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 394
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 395 ACAGCAACGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 454
QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 455 AGCGGATCTGGCCCTTACCAGATGTTCTGT 484

RESULT 4

CN419611 626 bp mRNA linear EST 16-MAY-2004
LOCUS 17000531390699 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN419611
ACCESSION CN419611.1 GI:47407205
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 626)
AUTHORS Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murgue,J., Fisk,G.J., Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R., Lebkowski,J and Stanton,L.W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation

JOURNAL COMMENT

Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com

Insert Length: 626 Std Error: 0.00.

FEATURES source

Location/Qualifiers
1..626
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/note="Oligo dT primed, full-length enriched cDNA library from embryoid body outgrowths derived from hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions."

ORIGIN

Alignment Scores:
Pred. No.: 6,9e-59 Length: 626
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x CN419611 (1-626)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 70 CCCATAGCTGAGAAGTGTTCCTGCTGGGACCTCTATGTCGGAGAACG 129
QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
DB 130 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTTGCTGCTGGGAGAGCAGC 189
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 190 GGACGCATCTGCTGCTAGAAATAGATGATCAACGATCAGGACACCAAGGACATCTAT 249
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
DB 250 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 309
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 310 ACAGCAACGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 369
QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 370 AGCGGATCTGGCCCTTACCAGATGTTCTGT 399

RESULT 5

CN4140993 737 bp mRNA linear EST 05-AUG-2002
LOCUS AU140993 PLACE4 Homo sapiens cDNA clone PLACE400626 5', mRNA
DEFINITION AU140993
ACCESSION AU140993
VERSION AU140993.1 GI:11002514
KEYWORDS EST.
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 737)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Iisgai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,

Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isogai,T.)
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES
source
1..737
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000626"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pME189FL3"

ORIGIN

Alignment Scores:
Pred. No.: 8.5e-59 Length: 737
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x AUI40993 (1-737)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 362 CCATAGCTGAGAGTGTTCATCATCTGCTGGGACTTCTATGTGTCGAGAAACG 421
QY 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
DB 422 TGGGAGAAGCCCTACCAAGCTGGATGATGTTAGATTGTACTTGCCTGGGAGAAGCGCAG 481
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 482 GGACGCATCATCTGCACCTTCAGAAATAGATGCAACGATCAGGACACACGACATCCCTAT 541
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 542 AGAATTGGAGACACTGGAGCAGAGAGGATATTCAGGAAACCTGCTCCAGTGCATCTGC 601
QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 602 ACAGGCAACGCCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 661
QY 101 SerGlySerGlyPropheThrAspValarg 110
DB 662 AGCGGATCTGGCCCTTCACCGATGTCGT 691

RESULT 6

BM715855
LOCUS BM715855 765 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-EJ0-ahj-h-11-0-UI.r2 UI-E-EJ0 Homo sapiens cDNA clone
UI-E-EJ0-ahj-h-11-0-UI 5', mRNA sequence.

ACCESSION BM715855

VERSION BM715855.1 GI:19029113

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 765)

REFERENCE

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE

Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE
97044477

PUBMED

COMMENT
889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers

1..765
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahj-h-11-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into p7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCAGCA; lens, CGATTAGCGA; eye anterior segment,
AATGCCGAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Alignment Scores:
Pred. No.: 8.91e-59 Length: 765
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x BM715855 (1-765)

QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 182 CCATAGCTGAGAGTGTTCATCATCTGCTGGGACTTCTATGTGTCGAGAAACG 241
QY 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
DB 242 TGGGAGAAGCCCTACCAAGCTGGATGATGTTAGATTGTACTTGCCTGGGAGAAGCGCAGC 301
QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 302 GGACGCATCATCTGCACCTTCAGAAATAGATGCAACGATCAGGACACACGACATCCCTAT 361
QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80


```
Db 362 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 421
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 422 ACAGCAACGCGCGAGGAGTGGAGAGTGTGAGAGGCACACCTCTGTGAGACCATCG 481
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 482 AGCGATCTGGCCCTTCACCGATGTTGCT 511

RESULT 7
AUI40971
LOCUS AUI40971 PLACE4 Homo sapiens cDNA clone IMAGE:303400583 5', mRNA EST 05-AUG-2002
DEFINITION AUI40971
ACCESSION AUI40971
VERSION AUI40971.1 GI:11002492
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 861)
AUTHORS Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,J., Sugiyama,T.,
Nishikawa,T., Nakamura,Y., Sugano,S., Masuho,Y. and Isogai,T.
TITLE HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,
Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sugano,S.,
Masuho,Y., Isogai,T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project: 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
LOCATION/Qualifiers
FEATURES
source
1..861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000583"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 1,04e-58 Length: 861
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x AUI40971 (1-861)

Qy 1 ProLeaAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 158 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGAGCTTCCTATGTGTCGAGAAACG 217
Qy 21 TrpGluLysProTyrClnGlyTrpMetValAspCysThrCysLeuGlyGluThrSer 40
Db 218 TGGGAGAGCCCTACCAAGCGTGGATGGATGGATGGATGGATGGATGGATGGATGGATGG 277
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 278 GGACGCATCACTTGGCACTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCTTAT 337
```

```
Qy 61 ArgIleCysAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 338 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC 397
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 398 ACAGCAACGCGCGAGGAGTGGAGAGTGTGAGAGGCACACCTCTGTGAGACCATCG 457
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 458 AGCGATCTGGCCCTTCACCGATGTTGCT 487

RESULT 8
BC078656
LOCUS BC078656
DEFINITION Homo sapiens cDNA clone IMAGE:30347017, containing frame-shift
errors.
ACCESSION BC078656
VERSION BC078656.1 GI:50925326
KEYWORDS HTc.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 7501)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buettow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heien,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Spapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Pahey,J.J., Helton,E., Kettelman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 7501)
AUTHORS Director MGC Project.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Stefan Hansson
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 168 Row: 0 Column: 20
This clone was selected for full length sequencing because it
```

passed the following selection criteria: matched mRNA gi: 16933543
This clone has the following problem: frame shifted.

```

FEATURES
source
Location/Qualifiers
1. .7501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30347017"
/tissue_type="Placenta, normal"
/clone_lib="NTH MGC_147"
/lab_host="DH10B"
/note="vector: pBluescript"

```

```

ORIGIN

Alignment Scores:

Pred. No.:          1.64e-57          Length:          7501
Score:             627.00           Matches:         110
Percent Similarity: 100.00%          Conservative:    0
Best Local Similarity: 100.00%        Mismatches:     0
Query Match:       100.00%           Indels:         0
DB:                3                Gaps:          0

US-09-940-235-4 COPY 150 259 (1-110) x BC078656 (1-7501)
/notes vector: pbasecript

```

QY		1	ProilealaGluLysCysPheAsnHisalalaGlyThrSerTyrValValGlyGluThr	20
DB		806	CCCATAGCTGAGAAGTGTTTGTATCATCTGCTGGACTTCCTATGTGTCGGAAACG	865
QY		21	TroGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer	40
DB		866	TGGGAGAAGCCCTACCAGGCTGGATGATGTGTAGATTGTACTTGCTCGGAGAACGCACG	925
QY		41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
DB		926	GGACGCATCACTTCGCATCTTAGAAAATAGATGCAACGATCAGGACACAGGACATCCCTAT	985
QY		61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
DB		986	AGAATTGGAGACACCTGGAGACAGAGAGTAAATCGAGGAACCTGCTCCAGTGCATCTGC	104
QY		81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
DB		1046	ACAGGCCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG	110
QY		101	SerGlySerGlyProPheThrAspValArg	110
DB		1106	AGCGGATCTGGCCCCCTTCAACGATGTTCCGT	1135

RESULT 9	734 bp	linear	EST 05-AUG-2002
AU140802	PLACE4 Homo sapiens	cdna clone	PLAC5400266 5', mRNA
AU140802	sequence.		
AU140802			
AU140802.1	GI:11002323		
EST.			
Homo sapiens (human)			
SOURCE			
KEYWORDS			
VERSION			
ACCESSION			
DEFINITION			
LOCUS			
AU140802			

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 734)
 Ota, T., Suzuki, Y., Saito, K., Saito, K., Yamamoto, J., Sugiyama, T.,
 Nishikawa, T., Nakamura, Y., Sugano, S., Masuho, Y. and Isoigai, T.
 HRI human cDNA project [Ota, T., Suzuki, Y., Saito, K., Ishii, S.,
 Yamamoto, J., Sugiyama, T., Nishikawa, T., Nakamura, Y., Sugano, S.,
 Masuho, Y., Isoigai, T.)
 Unpublished (2000)

CONTACT: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986

Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

```

FEATURES
source
nearx_research_institute
Location/Qualifiers
1. .734
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE4000266"
/tissue_type="placenta"
/clone_lib="PLACE4"
/note="Vector: pME18SFL3"

```

ORIGIN

Alignment Scores:		
Pred. No.:	5,018-58	Length:
Score:	620.00	Matches:
Percent Similarity:	99.09%	Conservative:
Best Local Similarity:	99.09%	Mismatches:
Query Match:	98.88%	Indels:
DB:	1	Gaps:

US-09-940-235-4 COPY 150 259 (1-110) x AUL40802 (1-734)

QY	1	Pro	Le	Ala	GIu	LyS	Cys	Phe	Asp	His	Ala	Ala	Gly	Thr	Ser	Tyr	Val	Val	GIu	Gly	Thr	20
Db	309	CC	CA	TAG	CTG	GAG	AGT	GTT	TG	AT	CA	TG	CTG	TG	GG	ACT	TCT	TAT	GTG	GTG	CG	368
QY	21	Trp	GIu	LyS	Pro	Tyr	Gln	Gly	Trp	Met	Val	Asp	Cys	Thr	Cys	Leu	Gly	GIu	GIu	Ser	40	
Db	369	TGG	GAG	AAG	CC	TAC	CA	AGG	TG	CA	TG	ATG	TG	TACT	TG	CTT	GG	CG	AAG	CG	428	
QY	41	Gly	Arg	Leu	Thr	Cys	Thr	Ser	Arg	Asn	Arg	Cys	Asn	Asp	Gln	Asp	Thr	Arg	Thr	Ser	60	
Db	429	GG	ACG	CAT	CAC	TTC	GC	ACT	TT	CAG	AAA	TAG	ATG	CA	AC	GAT	CAG	GAC	CA	CA	488	
QY	61	Arg	Gly	Gly	Asp	Thr	Trp	Ser	LyS	LyS	Asp	Asn	Arg	Gly	Val	Asn	Leu	Leu	Gln	Cys	80	
Db	489	AG	AA	TGG	AG	AC	ACT	TG	GAC	CA	AG	AG	AT	AT	CG	AG	AAA	CT	GT	CT	548	
QY	81	Thr	Gly	Asn	Gly	Arg	Gly	Glu	Trp	LyS	Cys	Glu	Arg	His	Thr	Ser	Val	Gln	Thr	Ser	100	
Db	549	AC	AG	CA	AC	CG	CG	AG	AG	AG	TG	GAA	GT	GT	GAG	GG	CA	CAC	TCT	CT	608	
QY	101	Ser	Gly	Ser	Gly	Pro	Phe	Thr	Asp	Val	Arg	110										
Db	609	AG	CG	GA	TCT	GG	CC	CTT	CAC	CG	AT	GTT	CGT	638								

RESULT 10

RESOLU 10
AI095589/c

LOCUS 103538370

FOCUS DEFINITION

NOT TYPED

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

1

REFERENCE

AUTHORS

TITLE

TO: MEMPHIS

**JOURNAL
CONTENT**

COMMENT

AI743013/c
 LOCUS w95a12.x1 Soares NSF P8 9W OT PA P SI linear EST 20-DEC-1999
 DEFINITION IMAGE:2371870 3' similar to gb:X02761_cds1 FIBRONECTIN PRECURSOR (HUMAN);, mRNA sequence.
 ACCESSION AI743013
 VERSION AI743013
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 560)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 This clone is available royalty-free through LILNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 827 Std Error: 0.00
 Seg primer: -40UP from Gibco
 High quality sequence stop: 453.
 FEATURES
 source
 1..560
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2371870"
 /lab_host="DH10B"
 /clone_lib="Soares NSF P8 9W OT PA P SI"
 /note="Organ: pooled; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and cloneIDs: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145047-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HF8-9W pool 1: 758280-760583, 772104-774407 Soares NBHPA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
 Pred. No.: 4.51e-57 Length: 560
 Score: 610.00 Matches: 107
 Percent Similarity: 97.27% Conservative: 0
 Best Local Similarity: 97.27% Mismatches: 3
 Query Match: 97.29% Indels: 0
 DB: 1 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x AI743013 (1-560)

Qy 1 ProlleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
 |||||
 Db 429 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTGGGACTTCCTATGTGTCGGAGAAACG 370
 |||||

Qy 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
 |||||
 Db 369 TGGAGAGAGCCCTACCAAGGCTGGATGATGTTGATTGCTTGGAGAGGCAGC 310
 |||||

Qy 41 GlyArgileThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
 |||||
 Db 309 GGACGCATCACTTGCACCTCTAGAAATAGATGATCAACGATCAGGACCCAAAGGGCATCCTAT 250
 |||||

Qy 61 ArgileGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCyslleCys 80
 |||||

Db 249 AGAATTGGAGACCCCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 190
 |||||

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
 |||||

Db 189 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 130
 |||||

Qy 101 SerGlySerGlyProPheThrAspValArg 110
 |||||

Db 129 AGCGGATCTGGCCCTTCACCGATGTTGCT 100
 |||||

RESULT 13
 CD613789/c
 LOCUS CD613789 548 bp mRNA linear EST 12-JAN-2004
 DEFINITION 56022208J1 FLP Homo sapiens cDNA, mRNA sequence.
 ACCESSION CD613789
 VERSION CD613789.1 GI:40262053
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 548)
 REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
 AUTHORS Circular rapid amplification of cDNA ends for high-throughput
 TITLE extension cloning of partial genes
 JOURNAL Genomics 84 (1), 205-210 (2004)
 COMMENT Contact: Fu GK
 Incyte Genomics, Inc.
 3160 Porter Dr., Palo Alto, CA 94304, USA
 Tel: 6508454102
 Email: gfu@incyte.com
 FEATURES
 source
 1..548
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="FLP"
 /note="Vector: pDrive Cloning Vector"

ORIGIN

Alignment Scores:
 Pred. No.: 1.56e-56 Length: 548
 Score: 605.00 Matches: 108
 Percent Similarity: 98.18% Conservative: 0
 Best Local Similarity: 98.18% Mismatches: 2
 Query Match: 96.49% Indels: 0
 DB: 6 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x CD613789 (1-548)

Qy 1 ProlleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
 |||||

Db 440 CCCATAGCTGAGAAGTGTGTTGATCATGCTGCTAGGACTTCCTATGTGTCGGAGAAACG 381
 |||||

Qy 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
 |||||

Db 380 TGGAGAGCCCTTACCAAGCCGGATGATGTTGATTGCTTGGGAGAGGCAGC 321
 |||||

Qy 41 GlyArgileThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
 |||||

Db 320 GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACACAGCATCCTAT 261
 |||||

Qy 61 ArgileGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCyslleCys 80
 |||||

Db 260 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 201
 |||||

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
 |||||

Db 200 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 141
 |||||

Qy 101 SerGlySerGlyProPheThrAspValArg 110
 |||||

Db 140 AGCGGATCTGGCCCTTACCGATGTTGCT 111

RESULT 14
CD613788
LOCUS CD613788 551 bp mRNA linear EST 12-JAN-2004
DEFINITION 56022208H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD613788
VERSION CD613788.1 GI:40262052
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 551)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.

FEATURES
Location/Qualifiers
1..551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"

ORIGIN
Alignment Scores:
Pred. No.: 1..578-56 Length: 551
Score: 605.00 Matches: 108
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 2
Query Match: 96.49% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x CD613788 (1-551)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 109 CCCATAGCTGAGAAGTGTTCATCATGCTGCTAGACCTCTCTATGTTGGTGGAGAAACG 168

Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
Db 169 TGGGAGAAGCCCTACCAAGCCCGGATGATGGTAGATTGTACTTGCCTGGGAGAAGCAGC 228

Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 229 GGACGCATCACTTGCACCTTCTAGAATAAGATGATGACGATCAGGACACCAAGGACATCCTAT 288

Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 289 AGAATGGAGACACCTGGAGCAGAGAGGATATCGAGGAACCTGCTCCAGTGCATCTGC 348

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 349 ACAGGCAACGCCGAGGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG 408

Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 409 AGCGGATCTGGCCCTTCACCGATGTTGCT 438

RESULT 15
BE009640
LOCUS BE009640 547 bp mRNA linear EST 05-JUN-2000
DEFINITION PMA-BN0172-100400-002-C11 BN0172 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE009640
VERSION BE009640.1 GI:8269873
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 547)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=PM4-BN0172-100400-002-C11&t3=2000-04-10&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 10
High quality sequence stop: 533.

FEATURES
Location/Qualifiers
1..547
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0172"
/notes="Organ: breast normal; Vector: puc18; Site: 1: Smal; Site: 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 3..34e-56 Length: 547
Score: 602.00 Matches: 105
Percent Similarity: 99.06% Conservative: 0
Best Local Similarity: 99.06% Mismatches: 1
Query Match: 96.01% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x BE009640 (1-547)

Qy 5 LysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThrTrpGluLysPro 24
Db 47 AAGTGTTCATCATGCTGCTGGACTTGTCTATGTTGGTGGAGAAACGTCGGAGAAGCCC 106

Qy 25 TyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySerGlyArgIleThr 44
Db 107 TACCAAGGCTGGATGATGTTAGATTGTACTTGCCTGGAGAAGCAGCAGCATCACT 166

Qy 45 CysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyrArgIleGlyAsp 64
Db 167 TGCACCTCTAGAATAAGATGATCAACAGATCAGACACAGGACATCCTATAGATTGGAGAC 226

Qy 65 ThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCysThrGlyAsnGly 84
Db 227 ACCTGGAGCAAGAAGGATATCGAGGAACCTGCTCCAGTGCATCTGCACAGCAACGGC 286

Qy 85 ArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSerSerGlySerGly 104
Db 287 CGAGGAGAGTGGAGTGTGAGAGGCCACACCTCTGTGCAGACCACACATCGAGCGGATCTGGC 346

Qy 105 ProPheThrAspValArg 110
Db 347 CCTTTCACCGATGTTTCGT 364

Search completed: November 6, 2004, 04:55:38
Job time : 1349.92 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	627	100.0	7679	4	US-09-220-133-38	Sequence 38, Appl
2	627	100.0	7680	4	US-09-023-655-1389	Sequence 1289, Ap
3	627	100.0	7680	5	PCR-US95-09819-6	Sequence 6, Appli
4	627	100.0	7803	2	US-08-551-356-1	Sequence 1, Appli
5	627	100.0	7803	5	PCR-US93-12687-1	Sequence 1, Appli
6	627	100.0	8044	4	US-09-566-921-135	Sequence 135, App
7	612	97.6	7705	1	US-08-259-569-16	Sequence 16, Appl
8	612	97.6	7705	2	US-08-826-885-16	Sequence 16, Appl
9	612	97.6	7705	6	5455158-2	Patent No. 5455158
10	213	34.0	567	1	US-08-142-449B-5	Sequence 5, Appli
11	194.5	31.0	986	1	US-07-637-250A-8	Sequence 8, Appli
12	194.5	31.0	986	1	US-08-145-061-8	Sequence 8, Appli

Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
Db 521 TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGCTGGGAGAGCAGC 580
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 581 GGACGCATCCTGCACCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 640
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGGAGCAAGAGGATAATCCAGGAAACCTGCTCCAGTGCATCTGC 700
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 760
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTCACCGATGTTGCT 790

RESULT 2

US-09-023-655-1289
; Sequence 1289, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; APPLICATION NUMBER:
; PRIORITY APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 931396
US-09-023-655-1289

Alignment Scores: 1.65e-67 Length: 7680
Pred. No.: 627.00 Matches: 110
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-940-235-4_COPY_150_259 (1-110) x US-09-023-655-1289 (1-7680)
Qy 1 ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 461 CCATAGCTGAGAGTGTGTTTGTATCATGCTGCTGGACATTCCTATGTGTCGGAGAAACG 520
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
Db 521 TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGCTGGGAGAGCAGC 580
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 581 GGACGCATCCTGCACCTTCTAGAAATAGATCAACGATCAGGACACAGGACATCCTAT 640
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGGAGCAAGAGGATAATCCAGGAAACCTGCTCCAGTGCATCTGC 700
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 760
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTCACCGATGTTGCT 790
RESULT 3
PCT-US95-09819-6
; Sequence 6, Application PC/TUS9509819
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: FIBRONECTIN FIBRIN-BINDING PEPTIDES, DNA
; TITLE OF INVENTION: CODING THEREFOR AND USES THEREOF
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09819
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/283,857
; FILING DATE: 01-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOLD-1A PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US95-09819-6

Alignment Scores:

Pred. No.: 1.65e-67 Length: 7680
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x PCT-US95-09819-6 (1-7680)

Qy 1 ProileAlaGluLysCysPheAspHisAlaGlyThrSerTyValValGlyGluThr 20
Db 461 CCATAGCTGAGAAGTGTCTTTCATCATGCTGCTGGACTTCCTATGTGTCGAGAAACG 520
Qy 21 TrpGluLysProTyGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
Db 521 TGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGCTGGGAGGAGCAGC 580
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 60
Db 581 GGACGATCACTTGCACCTCTAGAAATAGATGCAACGATCAGACACAGGACATCTAT 640
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 700
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrSer 100
Db 701 ACAGCAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 760
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTACCAGATGTTGCT 790

RESULT 4

US-08-551-356-1
Sequence 1, Application US/08551356
Patent No. 5830700
GENERAL INFORMATION:
APPLICANT: Irani, Meher
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/551,356
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/349,762
FILING DATE:
APPLICATION NUMBER: US/07/998,271
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Holly, Julie A
REGISTRATION NUMBER: 33-246
REFERENCE/DOCKET NUMBER: 92-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

LENGTH: 7803 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 6..7346
US-08-551-356-1

Alignment Scores:

Pred. No.: 1.69e-67 Length: 7803
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x US-08-551-356-1 (1-7803)

Qy 1 ProileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyValValGlyGluThr 20
Db 546 CCCATAGCTGAGAAGTGTCTTTCATCATGCTGCTGGACTTCCTATGTGTCGAGAAACG 605
Qy 21 TrpGluLysProTyGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
Db 606 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGCTGGGAGGAGCAGC 665
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTy 60
Db 666 GGACGATCACTTGCACCTCTAGAAATAGATGCAACGATCAGGACACAGGACATCTAT 725
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 726 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 785
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrSer 100
Db 786 ACAGCAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 845
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 846 AGCGGATCTGGCCCTTACCAGATGTTGCT 875

RESULT 5

PCT-US93-12687-1
Sequence 1, Application PC/TUS9312687
GENERAL INFORMATION:
APPLICANT: Irani, Meher H.
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: ZymoGenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12687
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/998,271
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 92-26PC

NUMBER OF SEQUENCES: 20
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/58,241
FILING DATE: 04-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 526,397
FILING DATE: 21-MAY-1990
APPLICATION NUMBER: 345,952
FILING DATE: 28-APR-1989
APPLICATION NUMBER: 231,951
FILING DATE: 29-DEC-1988
SEQ ID NO: 2:
LENGTH: 7705
5455158-2

Alignment Scores:
Pred. No.: 1,236-65 Length: 7705
Score: 612.00 Matches: 108
Percent Similarity: 98.18% Conservative: 0
Best Local Similarity: 98.18% Mismatches: 2
Query Match: 97.61% Indels: 0
DB: 6 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x 5455158-2 (1-7705)

QY 1 ProtleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 461 CCCATGCTGAGAAGTGTTCATGCTGCTGGACTCCCTATGTGTGCGAGAAACG 520

QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
DB 521 TGGGAGAGCCCTACCAAGCTGGATGATGTTAGTGTCTGCTGGAGAGGCAGC 580

QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 581 GGACGCATCACTTGCACCTCTAGAAATAGATGATCAACATGATCAGGACACACAGGACATCCTAT 640

QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 641 AGAATTTGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 700

QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
DB 701 ACAGCAACGCCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGAGACCATCG 760

QY 101 SerGlySerGlyProPheThrAspValArg 110
DB 761 AGCGATCTGGCCCTTCACCGATGTTCTGT 790

RESULT 10
US-08-142-449B-5
Sequence 5, Application US/08142449B
Patent No. 5668104
GENERAL INFORMATION:
APPLICANT: Nakahata, Tatsutoshi
APPLICANT: Kawano, Genji
APPLICANT: Sudo, Tetsuo
APPLICANT: Kojima, Katsuaki
TITLE OF INVENTION: Physiologically Active Protein and Hematopoietic Stem Cell Growth Agent
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nels T. Lippert, White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,449B
FILING DATE: 24-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lippert, Nels T.
REGISTRATION NUMBER: 25,888
REFERENCE/DOCKET NUMBER: 1145358-304
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-819-8582
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 567 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
US-08-142-449B-5

Alignment Scores:
Pred. No.: 1,76e-17 Length: 567
Score: 213.00 Matches: 35
Percent Similarity: 55.91% Conservative: 17
Best Local Similarity: 37.63% Mismatches: 39
Query Match: 33.97% Indels: 2
DB: 1 Gaps: 1

US-09-940-235-4_COPY_150_259 (1-110) x US-08-142-449B-5 (1-567)

QY 1 ProtleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
DB 7 CCTACGAGATCGCTGCTTGCACCCCTACACAGTTTCCATTCATGCGTGGAGATGAG 66

QY 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer 40
DB 67 TGGGACGAGTGTCTGAATCAGGCTTTAACTGTGTGCGAGTGTAGGCTTTGGAGT 126

QY 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
DB 127 GGTCAATTCAGATGTCATTCATGATGTCGATGCAATGGT-----GTGAACATAC 180

QY 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
DB 181 AAGATTGGAGAGAGTGGACCGCTCAGGGAGAAATGGCCAGATGATGAGCTGCATGT 240

QY 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHis 93
DB 241 CTGGGACGGAAGAGGAGGAGATTCAGTGTGACCCCTCAT 279

RESULT 11
US-07-637-250A-8
Sequence 8, Application US/07637250A
Patent No. 5342762
GENERAL INFORMATION:
APPLICANT: Mosher, Deane F.
APPLICANT: Sottile, Jane M.
TITLE OF INVENTION: FIBRONECTIN PURIFICATION VECTOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Carl R. Schwartz, Quarles & Brady
STREET: 411 East Wisconsin Avenue
CITY: Milwaukee
STATE: Wisconsin
COUNTRY: USA
ZIP: 53202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA: 435
;; APPLICATION NUMBER: US/07/637,250A
;; FILING DATE: 19910103
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schwartz, Carl R.
;; REGISTRATION NUMBER: 29,437
;; REFERENCE/DOCKET NUMBER: 96-296-9170-4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (414)277-5715
;; TELEFAX: (414)277-5591
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 986 base pairs
;; TYPE: NUCLEIC ACID
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Rattus norvegicus
US-07-637-250A⁸

Alignment Scores:
Pred. No.: 7.87e-15 Length: 986
Score: 194.50 Matches: 35
Percent Similarity: 62.82% Conservative: 14
Best Local Similarity: 44.87% Mismatches: 28
Query Match: 31.02% Indels: 1
DB: Gaps: 1

US-09-940-235-4_COPY_150_259 (1-110) x US-07-637-250A-8 (1-986)

Qy 14 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 33
Db 748 ACTTACACGTCACACGACGTTCCACAGCGTCACGAGGAGGACATATGCTGAACGTGT 807
Qy 34 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 53
Db 808 ACTGCTTCGGTCAGGCGCGGGCAGATGGAATGTGACCCCATCGACCATGCCAAGAT 867
Qy 54 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 73
Db 868 TCAGAGACCCGCGACATTTTACAGATTGTTGACTCCTGG--GAGAAGTTTGTGCATGTT 924
Qy 74 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluThrLysCysGlu 91
Db 925 GTCAGATACCACTGTTACTGTTACGCGCGGTGGCATTTGGGAGTGGCACTGCCAG 978

RESULT 12
US-08-145-061-8
; Sequence 8, Application US/08145061
; Patent No. 5460955
; GENERAL INFORMATION:
; APPLICANT: Mosher, Deane F.
; APPLICANT: Sottile, Jane M.
; TITLE OF INVENTION: FIBRONECTIN PURIFICATION VECTOR
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carl R. Schwartz, Quarles & Brady
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,061
; FILING DATE:

;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 07/637,250
;; FILING DATE: 03-JAN-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Schwartz, Carl R.
;; REGISTRATION NUMBER: 29,437
;; REFERENCE/DOCKET NUMBER: 96-296-9170-4
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (414)277-5715
;; TELEFAX: (414)277-5591
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 986 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; ORGANISM: Rattus norvegicus
US-08-145-061-8

Alignment Scores:
Pred. No.: 7.87e-15 Length: 986
Score: 194.50 Matches: 35
Percent Similarity: 62.82% Conservative: 14
Best Local Similarity: 44.87% Mismatches: 28
Query Match: 31.02% Indels: 1
DB: Gaps: 1

US-09-940-235-4_COPY_150_259 (1-110) x US-08-145-061-8 (1-986)

Qy 14 SerTyrValValGlyGluThrTrpGluLysProTyrGlnGlyTrpMetMetValAspCys 33
Db 748 ACTTACACGTCACACGACGTTCCACAGCGTCACGAGGAGGACATATGCTGAACGTGT 807
Qy 34 ThrCysLeuGlyGluGlySerGlyArgIleThrCysThrSerArgAsnArgCysAsnAsp 53
Db 808 ACTGCTTCGGTCAGGCGCGGGCAGATGGAATGTGACCCCATCGACCATGCCAAGAT 867
Qy 54 GlnAspThrArgThrSerTyrArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 73
Db 868 TCAGAGACCCGCGACATTTTACAGATTGTTGACTCCTGG--GAGAAGTTTGTGCATGTT 924
Qy 74 AsnLeuLeuGlnCysIleCysThrGlyAsnGlyArgGlyGluThrLysCysGlu 91
Db 925 GTCAGATACCACTGTTACTGTTACGCGCGGTGGCATTTGGGAGTGGCACTGCCAG 978

RESULT 13
US-09-404-879A-176/c
; Sequence 176, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; CURRENT FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 176
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(372)
; OTHER INFORMATION: n = A, T, C or G
US-09-404-879A-176

Alignment Scores:

Pred. No.: 1,64e-08 Length: 372
Score: 139.00 Matches: 25
Percent Similarity: 50.68% Conservative: 12
Best Local Similarity: 34.25% Mismatches: 34
Query Match: 22.17% Indels: 2
DB: 4 Gaps: 1

US-09-940-235-4_COPY_150_259 (1-110) x US-09-404-879A-176 (1-372)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
|||
::: |||
Db 234 CCTACGGATGACTCGTGGCTTTGACCCCTACACAGTTTCCCATTTATGCCGTTGGAGATGAG 175
|||
::: |||
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
|||
::: |||
Db 174 TGGGAACGAATCTCTGAATCAGGCTTTAAACTGTGTGCCAGTGTCTTAGGCTTTGGAAGT 115
|||
::: |||
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
|||
::: |||
Db 114 GGTCAATTCAGATGATGATTCATCTAGATGGTCCATGACAATGGT-----GTGAACCTAC 61
|||
::: |||
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 73
|||
::: |||
Db 60 AAGATTGGAGAGAGTGGGACCGTCAGGGAGAAAATGGA 22
|||
::: |||

RESULT 14

US-09-404-879A-218/c
; Sequence 218, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; APPLICANT: Algate, Paul A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.462C2
; CURRENT APPLICATION NUMBER: US/09/404,879A
; NUMBER OF SEQ ID NOS: 393
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 218
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-404-879A-218

Alignment Scores:
Pred. No.: 1,64e-08 Length: 372
Score: 139.00 Matches: 25
Percent Similarity: 50.68% Conservative: 12
Best Local Similarity: 34.25% Mismatches: 34
Query Match: 22.17% Indels: 2
DB: 4 Gaps: 1

US-09-940-235-4_COPY_150_259 (1-110) x US-09-404-879A-237 (1-372)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
|||
::: |||
Db 139 CCTACGGATGACTCGTGGCTTTGACCCCTACACAGTTTCCCATTTATGCCGTTGGAGATGAG 198
|||
::: |||
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
|||
::: |||
Db 199 TGGGAACGAATGCTCTGAATCAGGCTTTAAACTGTGTGCCAGTGTCTTAGGCTTTGGAAGT 258
|||
::: |||
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
|||
::: |||
Db 259 GGTCAATTCAGATGATGATTCATCTAGATGGTCCATGACAATGGT-----GTGAACCTAC 312
|||
::: |||
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 73
|||
::: |||
Db 313 AAGATTGGAGAGAGTGGGACCGTCAGGGAGAAAATGGA 351
|||
::: |||

Search completed: November 6, 2004, 04:59:16

Job time : 43.1644 secs

Alignment Scores:
Pred. No.: 1,64e-08 Length: 372
Score: 139.00 Matches: 25
Percent Similarity: 50.68% Conservative: 12
Best Local Similarity: 34.25% Mismatches: 34
Query Match: 22.17% Indels: 2
DB: 4 Gaps: 1

US-09-940-235-4_COPY_150_259 (1-110) x US-09-404-879A-218 (1-372)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
|||
::: |||
Db 234 CCTACGGATGACTCGTGGCTTTGACCCCTACACAGTTTCCCATTTATGCCGTTGGAGATGAG 175
|||
::: |||
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
|||
::: |||
Db 174 TGGGAACGAATCTCTGAATCAGGCTTTAAACTGTGTGCCAGTGTCTTAGGCTTTGGAAGT 115
|||
::: |||
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
|||
::: |||
Db 114 GGTCAATTCAGATGATGATTCATCTAGATGGTCCATGACAATGGT-----GTGAACCTAC 61
|||
::: |||
Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGly 73
|||
::: |||
Db 60 AAGATTGGAGAGAGTGGGACCGTCAGGGAGAAAATGGA 22
|||
::: |||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 6, 2004, 02:55:27 ; Search time 189.11 Seconds
(without alignments)
3131.797 Million cell updates/sec

Title: US-09-940-235-4_COPY_150_259

Perfect score: 627

Sequence: 1 PIAEKCFDHAAGTSYVGET.....ERHTSVQTTSGSGPFTDVR 110

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3611042 seqs, 2692057975 residues

Total number of hits satisfying chosen parameters: 7222084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09940235/runat_03112004_174041_11368/app_query.fasta_1.1045
-DB=Published Applications NA -QWIT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09940235@cgn_1_1034@runat_03112004_174041_11368
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DRV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
- 2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
- 18: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	627	100.0	777	10	US-09-940-235-3
2	627	100.0	1541	10	US-09-940-235-9
3	627	100.0	2096	10	US-09-940-235-12
4	627	100.0	2127	15	US-10-210-120-49
5	627	100.0	2443	17	US-10-741-601-70
6	627	100.0	2488	17	US-10-741-601-75
7	627	100.0	4295	15	US-10-144-194A-51
8	627	100.0	6510	17	US-10-741-601-72
9	627	100.0	6988	16	US-10-236-392-1
10	627	100.0	7361	16	US-10-236-392-3
11	627	100.0	7680	9	US-09-964-824A-574
12	627	100.0	7680	14	US-10-171-311-63
13	627	100.0	7680	15	US-10-236-031B-69
14	627	100.0	7680	15	US-10-374-979-75
15	627	100.0	7680	16	US-10-182-936A-75
16	627	100.0	7680	16	US-10-641-643-1289
17	627	100.0	7680	17	US-10-717-597-222
18	627	100.0	7680	18	US-10-788-792-79
19	627	100.0	7795	15	US-10-084-817-2
20	627	100.0	7823	17	US-10-741-601-77
21	627	100.0	7848	17	US-10-741-601-78
22	627	100.0	7867	13	US-10-098-841-6
23	627	100.0	7935	17	US-10-741-601-74
24	627	100.0	7959	17	US-10-741-601-81
25	627	100.0	8013	17	US-10-741-601-71
26	627	100.0	8027	16	US-10-447-161-8
27	627	100.0	8044	15	US-10-734-564-27
28	627	100.0	8044	15	US-10-240-965-121
29	627	100.0	8062	13	US-10-098-841-5
30	627	100.0	8137	13	US-10-098-841-8
31	627	100.0	8155	17	US-10-741-601-79
32	627	100.0	8226	17	US-10-741-601-69
33	627	100.0	8230	13	US-10-098-841-7
34	627	100.0	8278	17	US-10-741-601-82
35	627	100.0	8332	17	US-10-741-601-73
36	627	100.0	8371	17	US-10-741-601-76
37	612	97.6	7705	16	US-10-447-161-4
38	580.5	92.6	8329	9	US-09-917-800A-1731
39	580.5	92.6	8329	16	US-10-191-803-113
40	497.5	79.3	1782	10	US-09-940-235-11
41	397	63.3	927	14	US-10-198-846-12486
42	342	54.5	512	9	US-09-925-301-428
43	309.5	49.4	87467	17	US-10-741-601-5634
C 44	304.5	48.6	309	9	US-09-864-761-30472
C 45	294.5	47.0	1189	14	US-10-198-846-11556

ALIGNMENTS

RESULT 1
US-09-940-235-3
: Sequence 3, Application US/09940235
: Publication No. US20030059921A1
: GENERAL INFORMATION:
: APPLICANT: Kumar, Rajesh
: APPLICANT: Sahni, Girish
: APPLICANT: Roy, Chait
: APPLICANT: Rajagopal, Kammar
: APPLICANT: Nihalani, Deepak
: APPLICANT: Sundaram, Vasudha
: APPLICANT: Yadav, Mahavir
: TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
: TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
: TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
: TITLE OF INVENTION: PROTEIN
: FILE REFERENCE: 07064-009002
: CURRENT APPLICATION NUMBER: US/09/940,235
: CURRENT FILING DATE: 2002-04-09
: PRIOR APPLICATION NUMBER: 09/471,349
: PRIOR FILING DATE: 1999-12-23

```

; PRIOR APPLICATION NUMBER: IN 3825/DEL/98
; PRIOR FILING DATE: 1998-12-24
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3

```

Alignment Scores:	7.94e-72	Length:	777
Pred. No.:	627.00	Matches:	110
Score:		Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	10		

US-09-940-235-4 COPY 150 259 (1-110) x US-09-940-235-3 (1-777)

1	QY	Pro	le	ala	Glu	Lys	Cys	Phe	Asp	His	ala	ala	Gly	Thr	Ser	Tyr	Val	Val	Gly	Glu	Thr	20
448	DB	CCC	ATAG	CTG	AGA	AGT	GTT	TTT	GAT	CAT	GCT	GCT	GGG	ACT	TCT	CTAT	GTG	TCG	GAGA	AAC	G	507
21	QY	Trp	Glu	Lys	Pro	Tyr	Gln	Gly	Trp	Met	Val	Asp	Cys	Thr	Cys	Leu	Gly	Glu	Gly	Ser	40	
508	DB	TGG	GAGA	AGC	CC	TAC	CA	AGG	TGG	ATG	TGT	TACT	TG	CT	GGG	AGA	AGG	CAC	G	567		
41	QY	Gly	Arg	ile	Thr	Cys	Thr	Ser	Arg	Asn	Arg	Cys	Asn	Asp	Gln	Asp	Thr	Arg	Thr	Ser	Tyr	60
568	DB	GG	AGC	AT	CA	CT	CTG	AC	TT	C	TAG	AA	TAG	AT	CA	AG	CA	AGG	AC	AT	CC	627
61	QY	Arg	ile	Gly	Asp	Thr	Trp	Ser	Lys	Lys	Asp	Asn	Arg	Gly	Asn	Leu	Gln	Cys	ile	Cys	80	
628	DB	AGA	ATT	GG	AG	CA	CAC	CT	GG	AG	CA	AG	AG	AT	AT	CG	AG	AA	AC	CT	CT	687
81	QY	Thr	Gly	Asn	Gly	Arg	Gly	Glu	Trp	Lys	Cys	Glu	Arg	His	Thr	Ser	Val	Gln	Thr	Thr	Ser	100
688	DB	AC	AGG	CA	AC	GG	CC	GG	GAG	AGT	TGG	AA	GTG	TG	AG	AGG	CA	CA	CT	CT	CT	747
101	QY	Ser	Gly	Ser	Gly	Pro	Phe	Thr	Asp	Val	Arg	110										
748	DB	AG	CGG	AT	CT	GG	CC	CC	CTT	CAC	CG	AT	GTT	CGT	777							

RESULT 2

```

US-09-540-2335-9
/ Sequence 9, Application US/09940235
/ Publication NO. US20030059921A1
/ GENERAL INFORMATION:
/ APPLICANT: Kumar, Rajesh
/ APPLICANT: Sahni, Rajesh
/ APPLICANT: Roy, Chait
/ APPLICANT: Rajagopal, Kammarra
/ APPLICANT: Nihalani, Deepak
/ APPLICANT: Sundarha, Vasudha
/ APPLICANT: Yadav, Mahavir
/ TITLE OF INVENTION: NOVEL CLOT-SPEE
/ TITLE OF INVENTION: PROTEINS POSSI
/ TITLE OF INVENTION: CHARACTERISTI
/ TITLE OF INVENTION: PROTEIN
/ FILE REFERENCE: 07064-009002
/ CURRENT APPLICATION NUMBER: US/09/
/ CURRENT FILING DATE: 2002-04-09
/ PRIOR APPLICATION NUMBER: 09/471,3
/ PRIOR FILING DATE: 1998-12-23
/ PRIOR APPLICATION NUMBER: IN 3825/
/ PRIOR FILING DATE: 1998-12-24
/ NUMBER OF SEQ ID NOS: 28
/ SOFTWARE: FASTSEQ for Windows Vers
/ SEQ ID NO 9

```

```

; LENGTH: 1541
; TYPE: DNA
; ORGANISM: AR
; FEATURE:
; OTHER INFORM
US-09-940-235-9

```

Alignment Scores:	1.87e-71	Length:	1541
Pred. No.:	627.00	Matches:	110
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	100.00%	Gaps:	0
DB:	10		

US-09-940-235-4 COPY 150 259 (1-110) x US-09-940-235-9 (1-1541)

Qy	1	ProileAlaGluIysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	20
Db	1209	CCCATAGCTGAGAAGTGTGTTTGATCATCTGCTGGGACTTCCTATGTGGTCGGAGAAACG	1268
Qy	21	TriPGluIysProTyrGlnGlyTyrMetMetValAspCysThrCysIeuGlyGluGlySer	40
Db	1269	TGGGAGAAGCCCTACCAAGCTCGATGATGTAGATTGTACTTGCTGGGAGAAGCGCAGC	1328
Qy	41	GlyArgIleThrCysThrSerArgAsnAtqCysAsnAspGlnAspThrArgThrSerTyr	60
Db	1329	GGACGCATCACTTGCACCTCTAGAAATAGATGCAACGATCGGACACACAGGCATCTCAT	1388
Qy	61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Db	1389	AGAAATTGGAGACACCTGGACCAAGAGGATAATCGAGGAACCTGCTCCAGTGCATCTGC	1448
Qy	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db	1449	ACAGGCAACGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGTGCAGACCACATCG	1508
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Db	1509	AGCGGATCTGGCCCTTCCACGATGTTTCGT	1538

RESULT 3

US-09-940-235-12

Sequence 12, Application US/09940235
Publication No. US20030059921A1
GENERAL INFORMATION:
APPLICANT: Kumar, Rajesh
APPLICANT: Sahni, Girish
APPLICANT: Roy, Chait
APPLICANT: Rajagopal, Kammar
APPLICANT: Nihalani, Deepak
APPLICANT: Sundaram, Vasudha
APPLICANT: Yadav, Mahavir
TITLE OF INVENTION: NOVEL CLOT-SPECIFIC STREPTOKINASE
TITLE OF INVENTION: PROTEINS POSSESSING ALTERED PLASMINOGEN ACTIVATION
TITLE OF INVENTION: CHARACTERISTICS AND A PROCESS FOR THE PREPARATION OF SAID
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 07064-009002
CURRENT APPLICATION NUMBER: US/09/940,235
CURRENT FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 09/471,349
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: IN 3825/DEL/98
PRIOR FILING DATE: 1998-12-24
NUMBER OF SEQ ID NOS: 28
SOFTWARE: Fast-SEQ for Windows Version 4.0
SEQ ID NO 12
LENGTH: 2096
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hybrid cassette
US-09-940-235-12


```
Alignment Scores:
Pred. No.: 2.75e-71 Length: 2096
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x US-09-940-235-12 (1-2096)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 1764 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGACCTCCCTATGTGTCGAGAAACG 1823

Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
Db 1824 TGGGAGAACCCCTACCAAGCTGGATCATGTAGATTGCTTCCCTGGGAGAGGAGC 1883

Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 1884 GGACGCATCCTTGCCTCTAGAAATAGATGATCAACGATCAGGACACAGGACATCCTAT 1943

Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 1944 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 2003

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 2004 ACAGGCACGCGCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 2063

Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 2064 AGCGGATCTGGCCCTTACCAGATGTTTCGT 2093

RESULT 4
US-10-210-120-49
; Sequence 49, Application US/10210120
; Publication No. US20030175736A1
; GENERAL INFORMATION:
; APPLICANT: Chinnaiyan, Arul M.
; APPLICANT: Rubin, Mark A.
; APPLICANT: Sreekumar, Arun
; TITLE OF INVENTION: Expression Profile of Prostate Cancer
; FILE REFERENCE: UM-07221
; CURRENT APPLICATION NUMBER: US/10/210,120
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 60/309,581
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/334,468
; PRIOR FILING DATE: 2001-11-15
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 49
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-210-120-49

Alignment Scores:
Pred. No.: 2.8e-71 Length: 2127
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x US-10-210-120-49 (1-2127)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 597 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGACCTCCCTATGTGTCGAGAAACG 656
```

```
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
Db 657 TGGGAGAACCCCTACCAAGCTGGATCATGTAGATTGCTTCCCTGGGAGAGGAGCAGC 716

Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 717 GGACGCATCCTTGCCTCTAGAAATAGATGATCAACGATCAGGACACAGGACATCCTAT 776

Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 777 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 836

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 837 ACAGGCACGCGCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 896

Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 897 AGCGGATCTGGCCCTTACCAGATGTTTCGT 926

RESULT 5
US-10-741-601-70
; Sequence 70, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 2443
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-70

Alignment Scores:
Pred. No.: 3.32e-71 Length: 2443
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x US-10-741-601-70 (1-2443)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 913 CCCATAGCTGAGAAGTGTTCATCATGCTGCTGGGACTTCCTATGTGTCGAGAAACG 972

Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
Db 973 TGGGAGAACCCCTACCAAGCTGGATCATGTAGATTGCTTCCCTGGGAGAGGAGCAGC 1032

Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 1033 GGACGCATCCTTGCCTCTAGAAATAGATGATCAACGATCAGGACACAGGACATCCTAT 1092

Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 1153 ACAGGCACGCGCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1212

Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 1213 AGCGGATCTGGCCCTTACCAGATGTTTCGT 1242
```

RESULT 6

```

US-10-741-601-75
; Sequence 75, Application US/10741601
; Publication No. US20040186519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

```

Alignment Scores:	
Pred. No.:	3.4e-71
Score:	627.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	17
	Gaps: 0
	Indels: 0
	Mismatches: 0
	Conservative: 0
Matches:	110
Length:	2488

US-09-940-235-4 COPY 150 259 (1-110) x US-10-741-601-75 (1-2488)

Qy		1	ProIleAcladLysCysPheAsnHisAlaLaGlyThrSerTyrValValGlyGluThr	20
Db		913	CCCATAGCTGAGAAGTGTTTTGTATCATCTGCTGGACTTCCTATGTGGTCGGAGAAACG	972
Qy		21	TrpGluLysProTyrGlnGlyTyrMetMetValAspCysThrCysLeuGlyGluGlySer	40
Db		973	TGGGAGAAGCCCTACC AAGCGTGGATCATGTTAGATTGTACTTGCTTGGGAGAAGGCAGC	1032
Qy		41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Db		1033	GGACGCATCACTTGCACCTCTTAGAAAATAGATGC AACGATCAGACACAAGACATCTCAT	1092
Qy		61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Db		1093	AGAATTGGAGACACCTGGAGCAAGAAGGATAATCGAGAGAACCTGCTCCAGTGCATCTGC	1152
Qy		81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Db		1153	ACAGGCCAACGGCCGAGGAGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCACATCG	1212
Qy		101	SerGlySerGlyProPheThrAspValArg	110
Db		1213	AGCGGATCTGGCCCCCTTCAACCGATGTTTCGT	1242

RESULT 7

```

US-10-144-194A-51
; Sequence 51, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast
; FILE REFERENCE: 3U 103 R1
; Cancer Genes

```

Alignment Scores:		
Pred. No.:	6,73e-71	4295
Score:	627.00	110
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	15	Gaps: 0
US-09-940-235-4 COPY 150 259 (1-110) x US-10-144-194A-51 (1-4295)		

Qy	1	ProIleAlaGluLysCysPheAspHisAalaAclyThrSerTyrValValGlyGluThr	20
Dd	867	CCCATAGCTGAGAAGTGTTTTTGATCATGTGCTGGACTTCCTATGTGTCGGAGAAACG	926
Qy	21	TtpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer	40
Dd	927	TGGGAGAAGGCCATTACC AAGCGTGGATCATGTAGATTGTACTTCCTCGGGAGAAGGCAGC	986
Qy	41	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	60
Dd	987	GGACGCATCACTTGCATCTCTAGAAAATAGATGCAACGATCAGGACACAAGGACATCCTAT	1046
Qy	61	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys	80
Dd	1047	AGAAITGGAGACACCTGGAGCAGAGAGATAATCAGAGAAACCTGCTCCAGTGCATCTGC	1106
Qy	81	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	100
Dd	1107	ACAGCACACGCCGAGGAGAGTGGAAGTGTGAGAGGCACACCTCTGTGTCAGACCACATCG	1166
Qy	101	SerGlySerGlyProPheThrAspValArg	110
Dd	1167	AGCGGATCTGGCCCCCTTACCGATGTTTCGT	1196

RESULT 8

```

US-10-741-601-72
; Sequence 72, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01500
; CURRENT APPLICATION NUMBER: US/10741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 6510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-72

```

```

Alignment Scores:
Pred. No.:      1,138-70      Length:      6510
Score:          627.00      Matches:     110
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches:  0
Query Match:    100.00%      Indels:      0
DB:             17           Gaps:        0

US-09-940-235-4 COPY 150 259 (1-110) x US-10-741-601-72 (1-6510)

```

Oy		1	ProllealaglulysCysPheAspHisalaaclyThrSerTyrValvalGluThr	20
Dd	913		CCCATAGCTGAGAAGTGTTTTGTATCATCGTGGACTTCCTATGTGTCGGGAACG	972
Oy		21	TpdglulysProTyrlngInglyTrpMetMetValaspCystHrCysLeuglYgLuGlYSer	40
Dd	973		TGGGAGAAGGCCATTACC AAGGCTGGATGATGTAGATTGTACTTCCTCGGGAAGGCAGC	1032
Oy	41	GlvArgilethrCvsThrSerArqAsnArqCysAsnAspGlnAspThrArgThrSerTyr		60

Db 1033 GGACGCATCTGCTTCTAGAAATAGATGCAACGATCAGACACAAGGACATCCTAT 1092
Qy 61 ArgileGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 1093 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 1152
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 1153 ACAGGCAACGGCCGAGGAGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 1212
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 1213 AGCGGATCTGGCCCTTCACCGATGTTTCTG 1242

RESULT 9

US-10-236-392-1
; Sequence 1, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladamir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRochelle, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Pena, Carol A
; APPLICANT: Peyman, John A
; APPLICANT: Rastelli, Luca
; APPLICANT: Reiger, Daniel K
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Shenoy, Suresh
; APPLICANT: Shinkets, Richard A
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-442A
; CURRENT APPLICATION NUMBER: US/10/236,392
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US60/390,155
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION NUMBER: US09/635,949
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US60/318,765
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/357,303
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US60/367,753
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: US60/369,479
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: US09/659,634
; PRIOR FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: US60/318,120
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US60/318,130

; PRIOR FILING DATE: 2001-09-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 794
; SOFTWARE: Custom
; SEQ ID NO 1
; LENGTH: 6988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26)..(6986)
US-10-236-392-1
Alignment Scores:
Pred. No.: 1,24e-70 Length: 6988
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0
US-09-940-235-4_COPY_150_259 (1-110) x US-10-236-392-1 (1-6988)
Qy 1 ProileAlaGlyLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 461 CCCATAGCTGAGAAAGTGTTTTGATCATGCTGCTGGGACTTCTTATGTGTGGGAAACG 520
Qy 21 TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGluGlySer 40
Db 521 TGGGAGAAGCCCTACCAAGCTGGATGATGTAGATTGTACTTGCCTGGGAGAAGGCAGC 580
Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 581 GGACGCATCTTCTGCACTTCTAGAAATAGATGCAACGATCAGACACACAGACATCTCTAT 640
Qy 61 ArgileGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 700
Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGGCAACGGCCGAGGAGAGTGGAAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 760
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTCACCGATGTTTCTG 790

RESULT 10

US-10-236-392-3
; Sequence 3, Application US/10236392
; Publication No. US20040067490A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Burgess, Catherine, E
; APPLICANT: Casman, Stacie J
; APPLICANT: Catterton, Elina
; APPLICANT: Chapoval, Andrei
; APPLICANT: Crabtree, Julie
; APPLICANT: Edinger, Shlomit, R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Grosse, William M
; APPLICANT: Gusev, Vladamir
; APPLICANT: Kekuda, Ramesh
; APPLICANT: LaRochelle, William J
; APPLICANT: Li, Li
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Miller, Charles E
; APPLICANT: Millet, Isabelle
; APPLICANT: Padigar, Muralidhara

APPLICANT: Patturajan, Meera
APPLICANT: Pena, Carol A
APPLICANT: Peyman, John A
APPLICANT: Rastelli, Luca
APPLICANT: Reiger, Daniel K
APPLICANT: Rothenberg, Mark E
APPLICANT: Shenoy, Suresh
APPLICANT: Shimkets, Richard A
APPLICANT: Smithson, Glenda
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-442A
CURRENT APPLICATION NUMBER: US/10/236,392
CURRENT FILING DATE: 2002-09-06
PRIOR APPLICATION NUMBER: US09/540,763
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: US60/390,155
PRIOR FILING DATE: 2002-06-19
PRIOR APPLICATION NUMBER: US09/635,949
PRIOR FILING DATE: 2000-08-10
PRIOR APPLICATION NUMBER: US60/318,765
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: US60/357,303
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US60/367,753
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER: US60/369,479
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: US09/659,634
PRIOR FILING DATE: 2000-09-12
PRIOR APPLICATION NUMBER: US60/318,120
PRIOR FILING DATE: 2001-09-07
PRIOR APPLICATION NUMBER: US60/318,130
PRIOR FILING DATE: 2001-09-07
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 794
SOFTWARE: Custom
SEQ ID NO 3
LENGTH: 7361
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
FEATURE: (3)...(6663)
NAME/KEY: misc feature
LOCATION: (2447)...(2447)
OTHER INFORMATION: "n" = "a", "c", "t" or "g"
FEATURE:
NAME/KEY: misc feature
LOCATION: (3116)...(3116)
OTHER INFORMATION: "n" = "a", "c", "t" or "g"
FEATURE:
NAME/KEY: misc feature
LOCATION: (3161)...(3161)
OTHER INFORMATION: "n" = "a", "c", "t" or "g"
FEATURE:
NAME/KEY: misc feature
LOCATION: (5696)...(5696)
OTHER INFORMATION: "n" = "a", "c", "t" or "g"
US-10-236-392-3
Alignment Scores:
Pred. No.: 1.32e-70 Length: 7361
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0
US-09-940-235-4_COPY_150_259 (1-110) x US-10-236-392-3 (1-7361)
QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
|||||

Db 546 CCCATAGCTGAGAAAGTGTTCATCATCTGCTGGGACCTTCCTATGTCGCGAGAAACG 605
QY 21 TTPGLULYSPROTYRGINGLYTRPMTMETVALASPCTHRCYSLLEUGLYGLYSER 40
|||||
Db 606 TGGGAGAAGCCCTACCAGGCTGGATGATGGTAGATTGCTTGGGAGAAGGAGC 665
QY 41 GLYATGILETHRCYSTHRSERARGASNARGCYASNASPGINASPThrArgThrSerTyr 60
|||||
Db 666 GGAAGCATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 725
QY 61 ARGILEGLYASPTHRRPSTRYLVSASPASNARGLYASNLEULEUGLYCYLLECY 80
|||||
Db 726 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 785
QY 81 THRGLYASNGLYARGGLYGLUTRPLYSCYSGLUARGHISThrSerValGlnThrThrSer 100
|||||
Db 786 ACAGGCAACGGCGGAGGAGTGAAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 845
QY 101 SERGLYSERGLYPROPHETHRASPVALARG 110
|||||
Db 846 AGCGGATCTGGCCCTTCACCGATGTTGCT 875
RESULT 11
US-09-964-824A-574
; Sequence 574, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 574
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-574
Alignment Scores:
Pred. No.: 1.39e-70 Length: 7680
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-940-235-4_COPY_150_259 (1-110) x US-09-964-824A-574 (1-7680)
QY 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
|||||
Db 461 CCCATAGCTGAGAAAGTGTTCATCATCTGCTGGGACCTTCCTATGTCGCGAGAAACG 520
QY 21 TTPGLULYSPROTYRGINGLYTRPMTMETVALASPCTHRCYSLLEUGLYGLYSER 40
|||||
Db 521 TGGGAGAAGCCCTACCAGGCTGGATGATGGTAGATTGCTTGGGAGAAGGAGC 580
QY 41 GLYATGILETHRCYSTHRSERARGASNARGCYASNASPGINASPThrArgThrSerTyr 60
|||||
Db 581 GGACGCATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACAGGACATCCTAT 640
QY 61 ARGILEGLYASPTHRRPSTRYLVSASPASNARGLYASNLEULEUGLYCYLLECY 80
|||||
Db 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGCTCCAGTGCATCTGC 700

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 760
Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTACCAGATGTTTCGT 790

RESULT 12
US-10-171-311-63
; Sequence 63, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoerish, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-171-311-63

Alignment Scores:
Pred. No.: 1.39e-70 Length: 7680
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x US-10-171-311-63 (1-7680)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 461 CCCATAGCTGAGAAGTGTTCATCATGCTGTGGGACCTCTATGCTGGGAGAAACG 520

Qy 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
Db 521 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTTCTTCCCTGGGAGGAGCAGC 580

Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 581 GGACGCAATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACACGACATCCTAT 640

Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGTCTCCAGTGCATCTGC 700

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 760

Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTACCAGATGTTTCGT 790

RESULT 14
US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05

Db 761 AGCGGATCTGGCCCTTACCAGATGTTTCGT 790

RESULT 13
US-10-236-031B-69
; Sequence 69, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 69
; LENGTH: 7680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-69

Alignment Scores:
Pred. No.: 1.39e-70 Length: 7680
Score: 627.00 Matches: 110
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-940-235-4_COPY_150_259 (1-110) x US-10-236-031B-69 (1-7680)

Qy 1 ProfileAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr 20
Db 461 CCCATAGCTGAGAAGTGTTCATCATGCTGTGGGACCTCTATGCTGGGAGAAACG 520

Qy 21 TrpGluLysProTyrGlnGlyTrpMetValAspCysThrCysLeuGlyGlySer 40
Db 521 TGGGAGAAGCCCTACCAAGCTGGATGATGGTAGATTGTTCTTCCCTGGGAGGAGCAGC 580

Qy 41 GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr 60
Db 581 GGACGCAATCACTTGCACCTTCTAGAAATAGATGCAACGATCAGGACACACGACATCCTAT 640

Qy 61 ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuLeuGlnCysIleCys 80
Db 641 AGAATTGGAGACACCTGGAGCAAGAGGATAATCGAGAAACCTGTCTCCAGTGCATCTGC 700

Qy 81 ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer 100
Db 701 ACAGGCAACGGCCGAGGAGTGGAGTGTGAGAGGCACACCTCTGTGCAGACCATCG 760

Qy 101 SerGlySerGlyProPheThrAspValArg 110
Db 761 AGCGGATCTGGCCCTTACCAGATGTTTCGT 790

RESULT 14
US-10-374-979-75
; Sequence 75, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05

